communication protocol or extension of protocol, such as J2ME or SNMP may be used. In addition, terminal management operation of a wireless terminal may be performed in any other way, so that terminal management actions may be associated with conditions such as condition data. Other advantages will be recognized by those with ordinary skill in the art.

[00036] In one embodiment, the system and method described may be incorporated as an extension of the existing SyncML terminal management protocol in a non-invasive manner. For example, the extension information, such as the condition data, attribute data, and action data (i.e. wireless terminal based condition based control information) may be incorporated into the existing SyncML payload. These extensions to the existing SyncML payload are completely transparent to existing SyncML systems. As a result, systems capable of receiving the extension information will extract and interpret the extension information, while older systems will merely discard the extension information without error. Additionally, the extension to the existing SyncML protocol is extensible, and can therefore be easily customized or enhanced to address new requirements and criteria.

[00037] The system and method allows an operator to control the management functions of wireless mobile terminals individually on a per-terminal basis rather than by enforcing a management function on all terminals. Additionally, since the responsibility for terminal management execution and for the execution of multitasking multiple terminal management actions occurs within the wireless terminal 20 rather than in the terminal management network element 100, execution of a terminal management action is more likely to execute successfully. Since execution of the terminal management action on the terminal management network element 100 may be interrupted or aborted if the radio access network 22 becomes unavailable, processing of the terminal management actions by the wireless terminal 20 would be

independent of the radio access network 22. As a result, reliability is improved because the wireless terminal executes the terminal management actions independently from the terminal management network element 100, even if access to the radio access network 16 is unavailable (i.e. in an area without signal).

[00038] Another advantage includes the ability of the wireless terminal 22 to be self-adaptive to reflect changes on the specific terminal. Other advantages include a reduction in bandwidth for performing terminal management actions on the wireless terminal 20 rather than on the terminal management network element 100. As a result, high data rate communication over the radio access network 22 for performing the terminal management operation is not required, and therefore, a reduction in communication bandwidth is realized. Consequently, the cost and effort associated with wireless terminal management is reduced.

[00039] It should be understood that the implementation of other variations and modifications of this invention and its various aspects will be apparent to those of ordinary skill in the art, and that the invention is not limited by specific embodiments described. It is therefore contemplated to cover by the present invention, any and all modifications, variations, or equivalents that fall within the spirit and scope of the basic underlying principles disclosed and claimed herein.

WE CLAIM:

1. A method in a wireless terminal comprising:

determining that a condition associated with at least one terminal management action is satisfied based on extensible markup language code that includes at least wireless terminal condition based control information that includes at least terminal condition information and action data wherein the action data identifies the at least one terminal management action associated with the wireless terminal; and

executing conditionally, by the wireless terminal, the terminal management action identified by the received action data.

2. The method of claim 1 wherein determining that the condition associated with at least one terminal management action is satisfied based on extensible markup language code that includes at least wireless terminal condition based control information includes: attribute data associated with the action data; and condition data associated with the attribute data.

3. The method of claim 2 comprising:

obtaining current terminal attribute values in response to the received attribute data at the wireless terminal; and

comparing the obtained current terminal attribute values to the condition data to determine if the condition data is met.

4. The method of claim 2 comprising:

scheduling a current terminal management action to conditionally execute according to whether a previous terminal management action is completed; and

conditionally executing the current terminal management action if the condition data associated with the current terminal management action is met wherein the condition data requires that the previous terminal management action is executed.

5. The method of claim 1 comprising:

obtaining wirelessly received wireless terminal based control information including a plurality of action data corresponding to a plurality of terminal management actions, a plurality of corresponding attribute data associated with the corresponding plurality of action data, a plurality of corresponding condition data associated with the corresponding plurality of attribute data, and

obtaining a plurality of current terminal attribute values in response to obtaining the plurality of corresponding attribute data associated with the corresponding plurality of action data;

comparing the plurality of current terminal attribute values to the corresponding condition data to determine if the condition data is met; and

conditionally executing the plurality of terminal management actions if the condition data associated with the corresponding terminal management action is satisfied wherein the condition data requires that a previous terminal management action is executed.

6. The method of claim 1 wherein, the terminal condition information associated with the terminal management action includes data representing at least one of: a wireless terminal state, a network state, a duration since a last terminal management action occurred, a number of action repetitions, a completion of a terminal management action, a software upgrade, a time of day, a terminal management action schedule, a wireless application protocol (WAP) configuration, a user setting, a signal level, a battery level, and terminal location coordinates.

7. The method of claim 1 where the wireless terminal condition based control information includes at least one of an, assert command, schedule command, and perform command, wherein:

the assert command associates the terminal condition information with the received action data and with exception action data if the condition data is not satisfied,

the schedule command associates the terminal condition information with data including at least one of a time of day, a time interval between successive executions of the terminal management action, number of repetitions to execute the terminal management action, repetition rate of executing the terminal management action, monitoring, measuring and reporting an elapsed period of time to perform a terminal management action, and a first run of a scheduled execution of a terminal management action, and

the perform command associates action data with an attribute value including at least one of a network attribute value, and a wireless mobile terminal attribute value.

8. A wireless terminal comprising:

a wireless transceiver that receives extensible markup language code that includes at least wireless terminal condition based control information that includes at least terminal condition information and action data wherein the action data identifies at least one terminal management action associated with the wireless mobile terminal; and

a wireless terminal condition based controller, coupled to the wireless transceiver,

the wireless terminal condition based controller for:

determining whether a condition associated with the terminal management action is satisfied, based on the terminal condition information, and

executing, the terminal management action identified by the action data if the condition associated with the terminal management action is satisfied.

9. The wireless terminal condition based controller of claim 8 comprising:

a command parser coupled to the wireless transceiver to parse the wireless terminal condition based control information to identify the terminal condition information and the action data associated with the terminal management action.

10. The wireless terminal of claim 9 wherein the wireless terminal condition based controller comprises:

a policy reader coupled to the command parser to read the action data, attribute data, and condition data associated with the terminal management action;

an attribute monitor coupled to the policy reader to receive attribute data from the policy reader and to obtain current terminal attribute values in response to the received attribute data; and

a task manager coupled to the policy reader and to the attribute monitor to compare the current terminal attribute values from the attribute monitor, with the received condition data to determine if the condition data is satisfied.

11. The wireless terminal of claim 10 wherein the task manager:

schedules a current terminal management action to execute according to whether a previous terminal management action is completed; and

conditionally executes the current terminal management action in accordance with the condition data associated with the current terminal management action, wherein the condition data requires that a previous terminal management action is executed.

12. The wireless terminal of claim 8 wherein the wireless terminal condition based controller comprises:

a processing device;

memory containing executable instructions that causes the processing devices to:

identify the at least wireless terminal condition based control information that includes the at least terminal condition information and the

action data wherein the action data identifies at least one terminal management action associated with the wireless mobile terminal; and

in response to determining that the condition associated with the terminal management action is satisfied based on the terminal condition information, conditionally execute, the terminal management action identified by the action data.

13. The wireless terminal of claim 8 wherein the terminal condition information associated with the terminal management action is determined at an event selected from the group consisting of a polling event, an interrupt event, a predetermined time of day, and a predetermined period of elapsed time.

14. The wireless terminal of claim 8 where the extensible markup language code includes at least one of an assert command, a schedule command, and a perform command, wherein:

the assert command associates the terminal condition information with the received action data and associates the terminal condition information with exception action data if the condition data is not satisfied,

the schedule command associates the terminal condition information with data including at least one of a time of day, time interval between successive executions of the terminal management action, number of repetitions to execute the terminal management action, repetition rate of executing the terminal management action, monitoring, measuring and reporting an elapsed period of time to execute the terminal

management action, and a first run of a scheduled execution of a terminal management action, and

the perform command associates the action data with an attribute value including at least one of a network attribute value and an attribute value associated with the wireless mobile terminal.

15. The wireless terminal of claim 8 wherein the wireless terminal based control information identifies a plurality of terminal management actions, wherein the wireless terminal executes the plurality of terminal management actions in accordance with the corresponding terminal condition information wherein the terminal condition information requires that a previous terminal management action is executed.

16. A storage medium having stored therein instructions executable by a processor that causes the processor to:

receive, extensible markup language code that includes at least wireless terminal condition based control information that includes at least terminal condition information and action data wherein the action data identifies at least one terminal management action associated with the wireless mobile terminal; and

in response to determining that a condition associated with the terminal management action is satisfied based on the terminal condition information, conditionally execute, by the processor, the terminal management action identified by the action data.

17. A wireless terminal comprising:

a wireless transceiver that receives extensible mark-up language code that includes a payload, including at least wireless terminal condition based control information; and

a wireless terminal condition based controller comprising:

a command parser coupled to the wireless transceiver that parses the wireless terminal condition based control information wherein the terminal condition information comprises:

action data associated with at least one terminal management action;

attribute data associated with the action data; and

condition data associated with the attribute data,

a policy reader coupled to the command parser operative to read the action data, attribute data and condition data associated with the least one terminal management action;

an attribute monitor coupled to the policy reader to receive the attribute data from the policy reader and to locally obtain current terminal attribute values in response to the received attribute data; and

a task manager coupled to the policy reader to compare the current terminal attribute values to the received condition data to:

determine if the condition data is met, and

conditionally execute the terminal management action identified by the action data.

18. A terminal management network element comprising:

a wireless terminal condition based control information generator that generates extensible markup language code that includes wireless terminal condition based control information that includes:

selected action data based on selectable action data associated with a terminal management action for at least one wireless mobile terminal;

selected condition data based on selectable condition data wherein the selected condition data is associated with the selectable action data; and

selected attribute data based on selectable attribute data wherein the selected attribute data is associated with the selectable action data,

wherein the wireless terminal condition based control information is generated according to the selected action data, the selected condition data and the selected attribute data; and

a transceiver coupled to the wireless terminal control based control information generator to transmit the wireless terminal control based information.

19. The terminal management network element of claim 18 wherein the wireless terminal condition based control information generator generates a plurality of terminal management actions defined by a plurality of action data, and

wherein the terminal management network element defines condition data associated with a first terminal management action to conditionally execute according to whether a second terminal management action completes execution.

20. The terminal management network element of claim 18 wherein the wireless management terminal network element generator:

generates wireless terminal condition based control information to conditionally execute a plurality of terminal management actions each associated with corresponding selected condition data on a plurality of wireless mobile terminals, and

sends network attribute values to the plurality of wireless mobile terminals to compare the network attribute values with the selected condition data to determine if the selected condition data is met.

21. A method for managing a terminal management network element comprising:

presenting selectable action data associated with a terminal management action;

receiving selected action data in response to presenting the selectable action data;

presenting selectable condition data associated with the selectable action data wherein the selected condition data is associated with the selected action data;

receiving selected condition data in response to presenting the selectable condition data;

presenting selectable attribute data associated with the selectable action data wherein the selected attribute data is associated with the selected action data;

receiving selected attribute data in response to presenting the selectable attribute data, and

generating wireless terminal condition based control information according to the selected action data, the selected condition data, and the selected attribute data wherein the wireless terminal condition based control information is represented by extensible markup language code for at least one wireless mobile terminal.

22. The method of claim 21 wherein the selected action data identifies a plurality of terminal management actions, the method comprising;

associating selected condition data with a first terminal management action to conditionally execute the first terminal management actions on a single wireless terminal according to whether a second terminal management action is completed; and

generating the plurality of current terminal management actions associated with the selected condition data.

23. The method of claim 21 comprising;

generating wireless terminal condition based control information for a plurality of wireless mobile terminals; and

sending network attribute values to the plurality of wireless mobile terminals to allow comparing the network attribute values with the selected condition data by each of the plurality of wireless mobile terminals to determine if the selected condition data is met.

24. A terminal management system comprising:

a terminal management network element comprising:

a wireless terminal condition based control information generator that generates wireless terminal condition based control information comprising:

selected action data based on selectable action data associated with a terminal management action to execute on at least one wireless mobile terminal;

selected condition data based on selectable condition data wherein the selected condition data is associated with the selected action data;

selected attribute data based on selectable attribute data wherein the selected attribute data is associated with the selected action data,

wherein the wireless terminal condition based control information is generated according to the selected action data, the selected condition data and the selected attribute data, wherein the wireless terminal condition based control information is represented by extensible markup language code;

at least one wireless terminal comprising:

a wireless transceiver that receives the at least wireless terminal condition based control information, the selected attribute data, the selected condition data, the selected action data and the terminal management action associated with the wireless mobile terminal; and

a wireless terminal condition based controller, operatively coupled to the wireless transceiver, and operative to

determine that the selected condition data associated with the terminal management action is satisfied, and

conditionally execute the terminal management action identified by the selected action data if the selected condition data associated with the terminal management action is satisfied.

25. The terminal management system of claim 24 wherein the wireless terminal condition based control information generator generates a plurality of terminal management actions defined by the selected action data and the associated selected condition data, and defines the selected condition data associated with each terminal management action to conditionally execute in accordance with the selected condition data associated with each respective terminal management action.

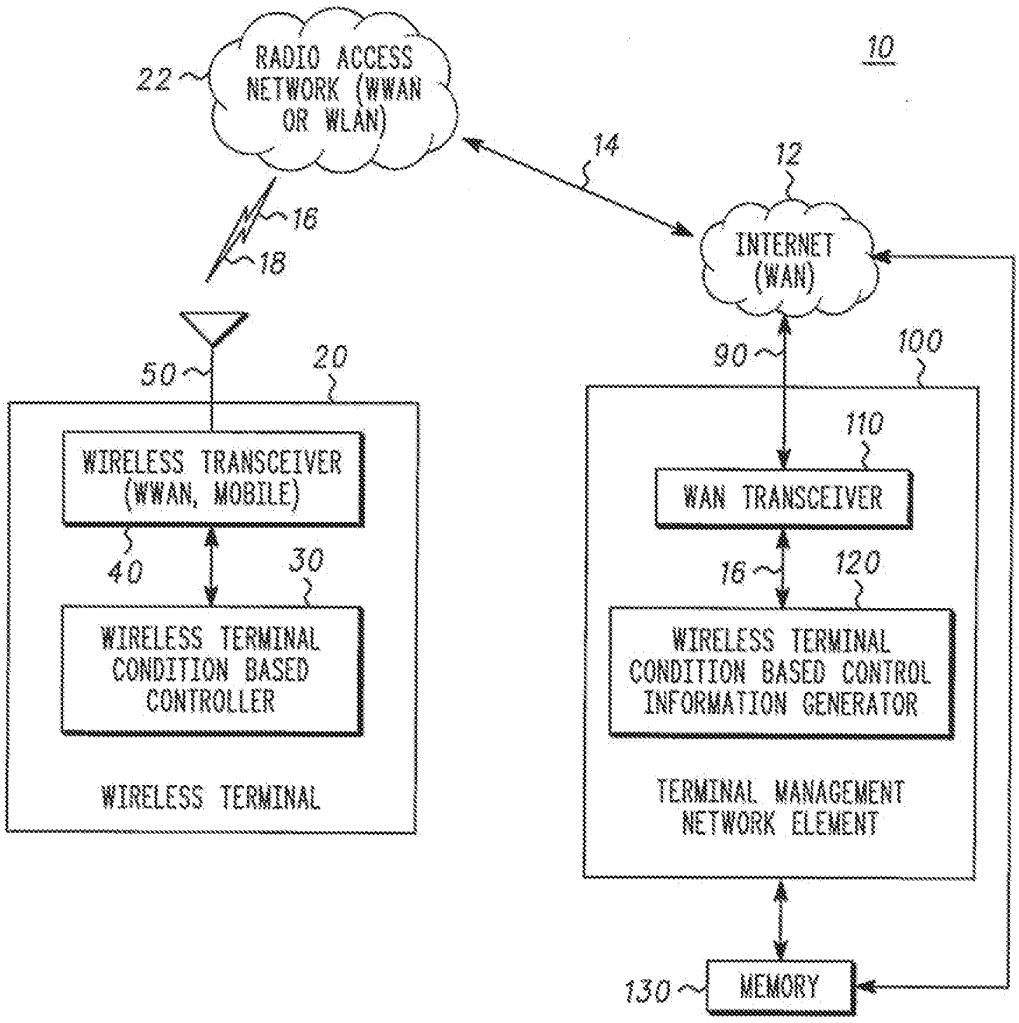

RADIO ACCESS NETWORK (WWAN OR WLAN)
22
10
14
12
16
18
INTERNET (WAN)
50
20
90
100
WIRELESS TRANSCEIVER (WWAN, MOBILE)
110
WAN TRANSCEIVER
40
30
16
120
WIRELESS TERMINAL CONDITION BASED CONTROLLER
WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION GENERATOR
WIRELESS TERMINAL
TERMINAL MANAGEMENT NETWORK ELEMENT
130
MEMORY

FIG. 1

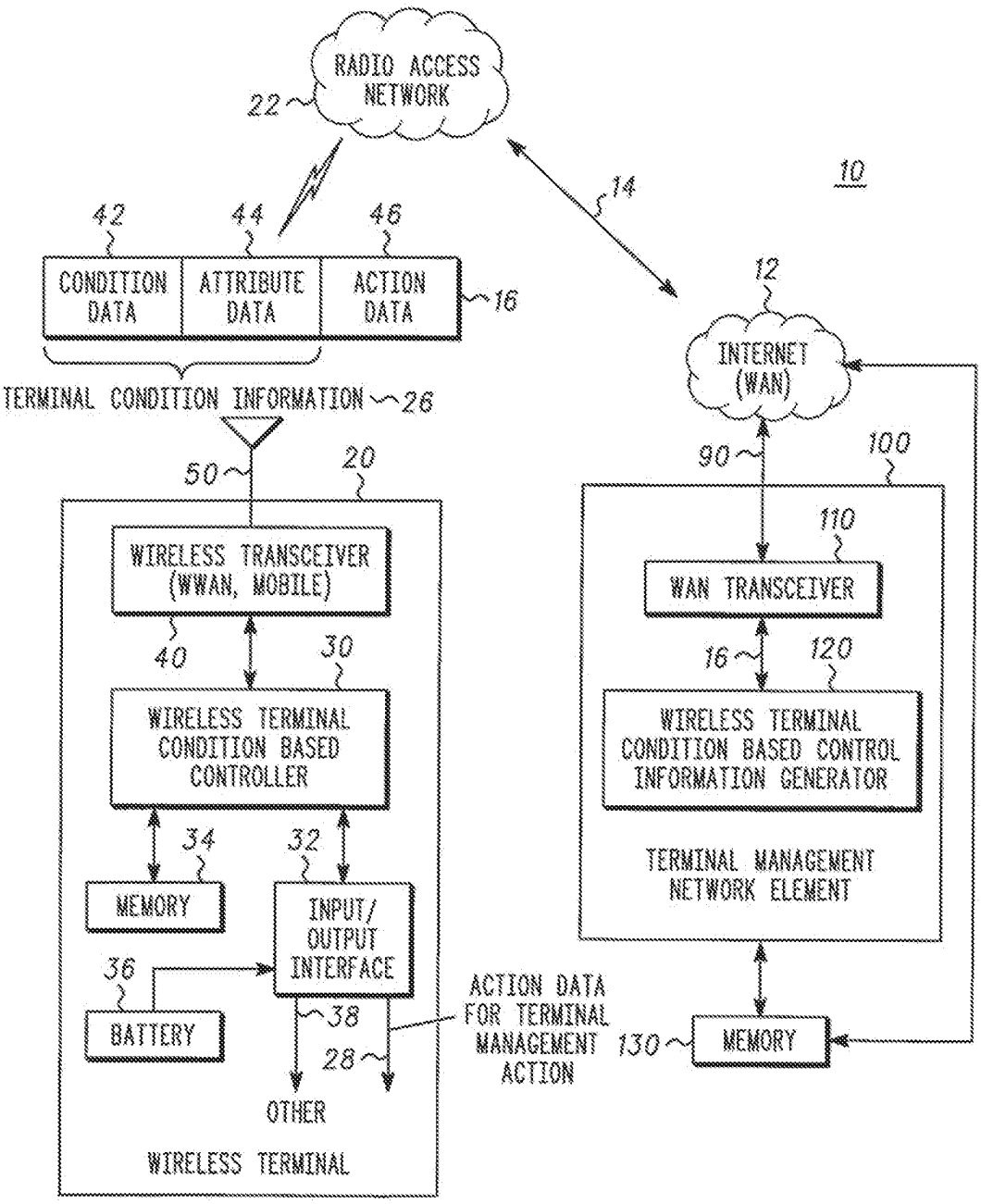

RADIO ACCESS NETWORK
22
14
10
42
44
46
CONDITION DATA
ATTRIBUTE DATA
ACTION DATA
16
12
INTERNET (WAN)
TERMINAL CONDITION INFORMATION
26
50
20
90
100
WIRELESS TRANSCEIVER (WWAN, MOBILE)
110
WAN TRANSCEIVER
40
30
16
120
WIRELESS TERMINAL CONDITION BASED CONTROLLER
WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION GENERATOR
34
32
MEMORY
INPUT/ OUTPUT INTERFACE
TERMINAL MANAGEMENT NETWORK ELEMENT
36
ACTION DATA FOR TERMINAL MANAGEMENT ACTION
BATTERY
38
28
130
MEMORY
OTHER
WIRELESS TERMINAL

FIG. 2

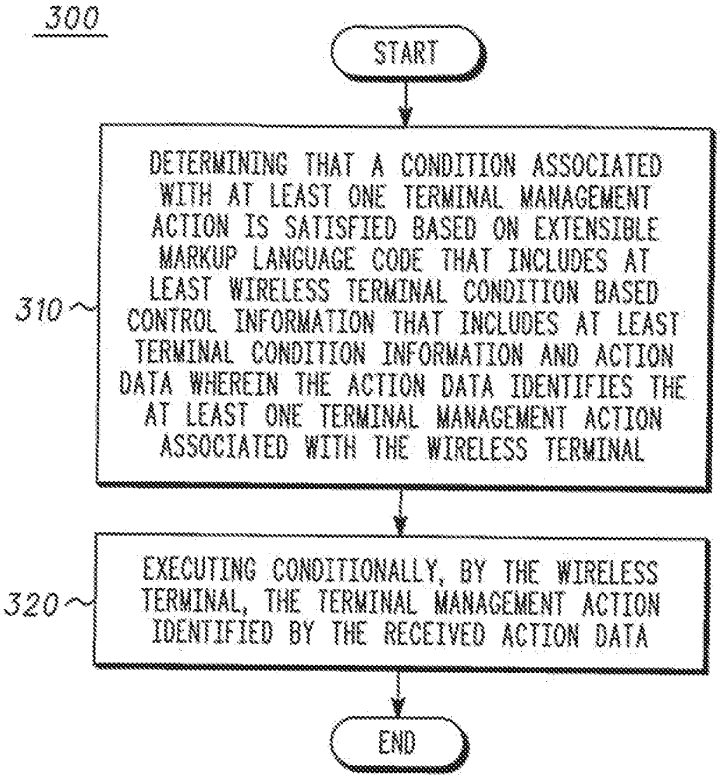

300
START
DETERMINING THAT A CONDITION ASSOCIATED WITH AT LEAST ONE TERMINAL MANAGEMENT ACTION IS SATISFIED BASED ON EXTENSIBLE MARKUP LANGUAGE CODE THAT INCLUDES AT LEAST WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION THAT INCLUDES AT LEAST TERMINAL CONDITION INFORMATION AND ACTION DATA WHEREIN THE ACTION DATA IDENTIFIES THE AT LEAST ONE TERMINAL MANAGEMENT ACTION ASSOCIATED WITH THE WIRELESS TERMINAL
310
EXECUTING CONDITIONALLY, BY THE WIRELESS TERMINAL, THE TERMINAL MANAGEMENT ACTION IDENTIFIED BY THE RECEIVED ACTION DATA
320
END

FIG. 3

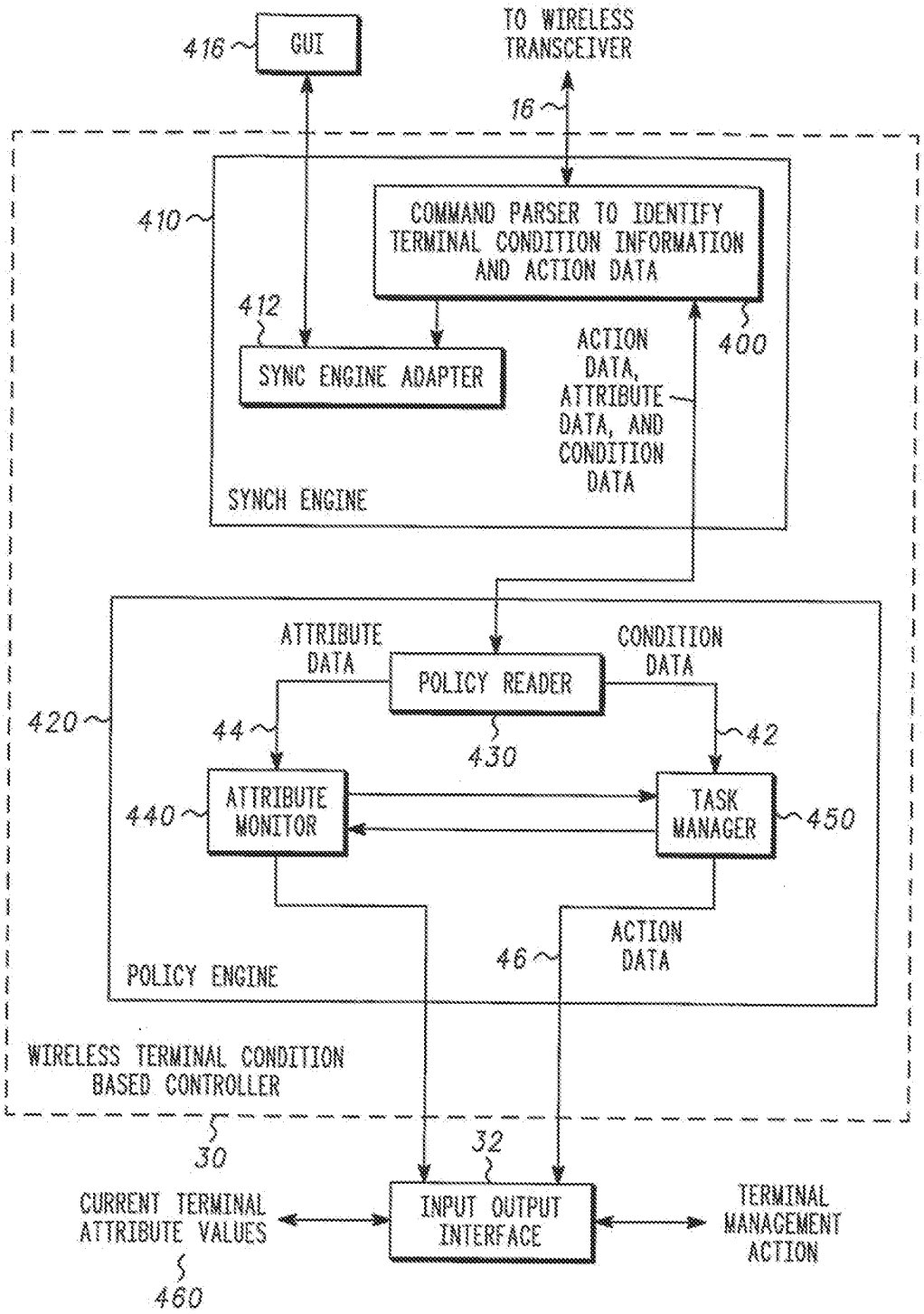

416
GUI
TO WIRELESS TRANSCEIVER
16
410
COMMAND PARSER TO IDENTIFY TERMINAL CONDITION INFORMATION AND ACTION DATA
412
400
SYNC ENGINE ADAPTER
ACTION DATA, ATTRIBUTE DATA, AND CONDITION DATA
SYNCH ENGINE
ATTRIBUTE DATA
POLICY READER
CONDITION DATA
420
44
42
430
440
ATTRIBUTE MONITOR
TASK MANAGER
450
46
ACTION DATA
POLICY ENGINE
WIRELESS TERMINAL CONDITION BASED CONTROLLER
30
32
CURRENT TERMINAL ATTRIBUTE VALUES
460
INPUT OUTPUT INTERFACE
TERMINAL MANAGEMENT ACTION

FIG. 4

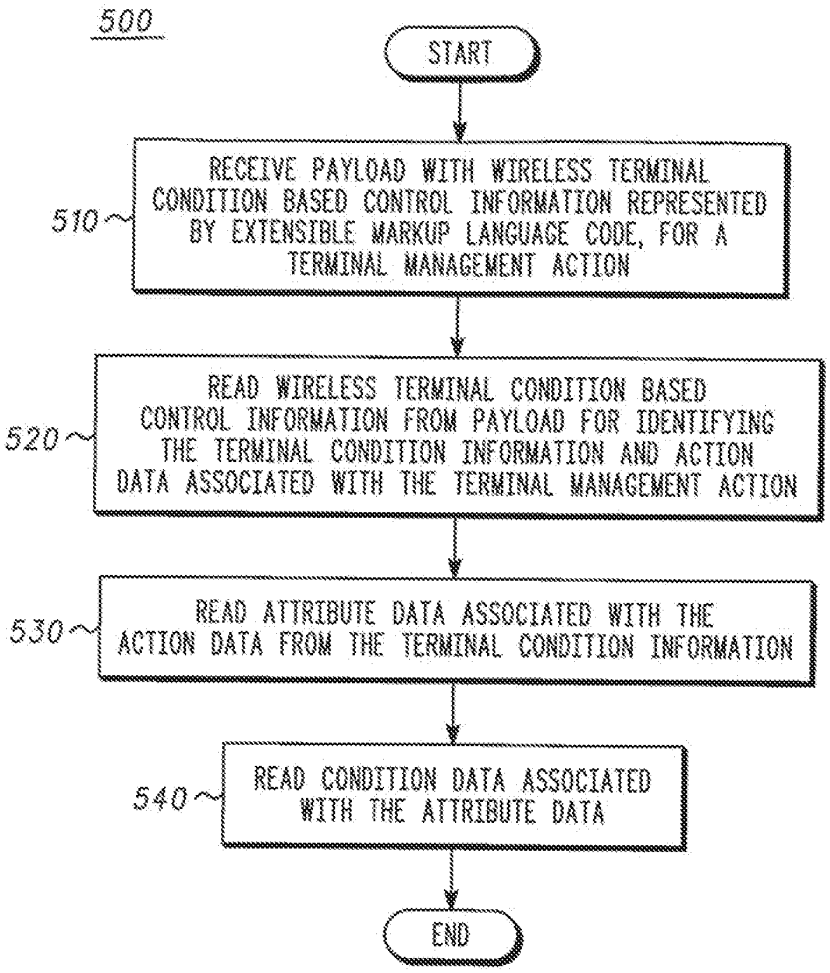

500
START
510
RECEIVE PAYLOAD WITH WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION REPRESENTED BY EXTENSIBLE MARKUP LANGUAGE CODE, FOR A TERMINAL MANAGEMENT ACTION
520
READ WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION FROM PAYLOAD FOR IDENTIFYING THE TERMINAL CONDITION INFORMATION AND ACTION DATA ASSOCIATED WITH THE TERMINAL MANAGEMENT ACTION
530
READ ATTRIBUTE DATA ASSOCIATED WITH THE ACTION DATA FROM THE TERMINAL CONDITION INFORMATION
540
READ CONDITION DATA ASSOCIATED WITH THE ATTRIBUTE DATA
END

FIG. 5

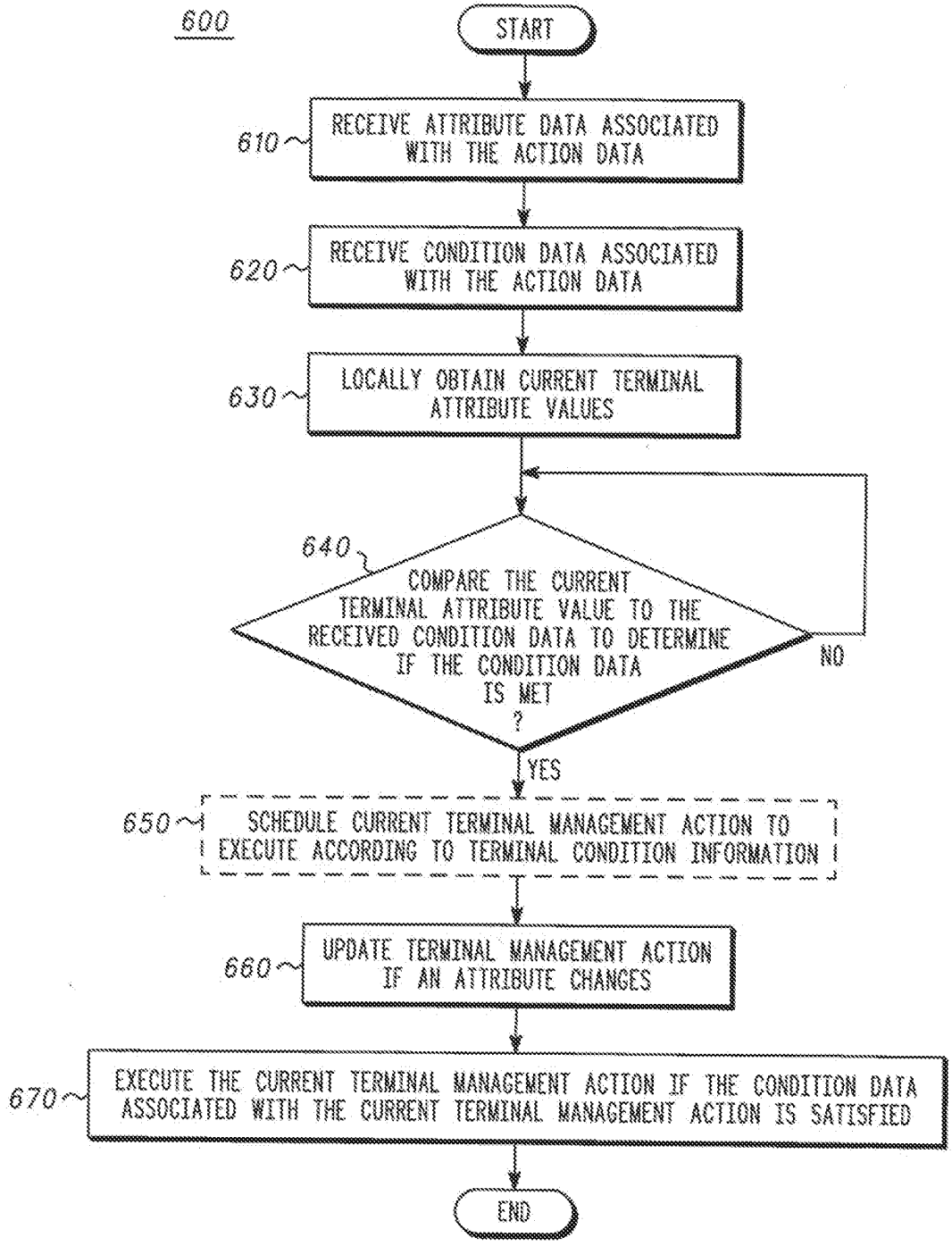

600
START
610
RECEIVE ATTRIBUTE DATA ASSOCIATED WITH THE ACTION DATA
620
RECEIVE CONDITION DATA ASSOCIATED WITH THE ACTION DATA
630
LOCALLY OBTAIN CURRENT TERMINAL ATTRIBUTE VALUES
640
COMPARE THE CURRENT TERMINAL ATTRIBUTE VALUE TO THE RECEIVED CONDITION DATA TO DETERMINE IF THE CONDITION DATA IS MET ?
NO
YES
650
SCHEDULE CURRENT TERMINAL MANAGEMENT ACTION TO EXECUTE ACCORDING TO TERMINAL CONDITION INFORMATION
660
UPDATE TERMINAL MANAGEMENT ACTION IF AN ATTRIBUTE CHANGES
670
EXECUTE THE CURRENT TERMINAL MANAGEMENT ACTION IF THE CONDITION DATA ASSOCIATED WITH THE CURRENT TERMINAL MANAGEMENT ACTION IS SATISFIED
END

FIG. 6

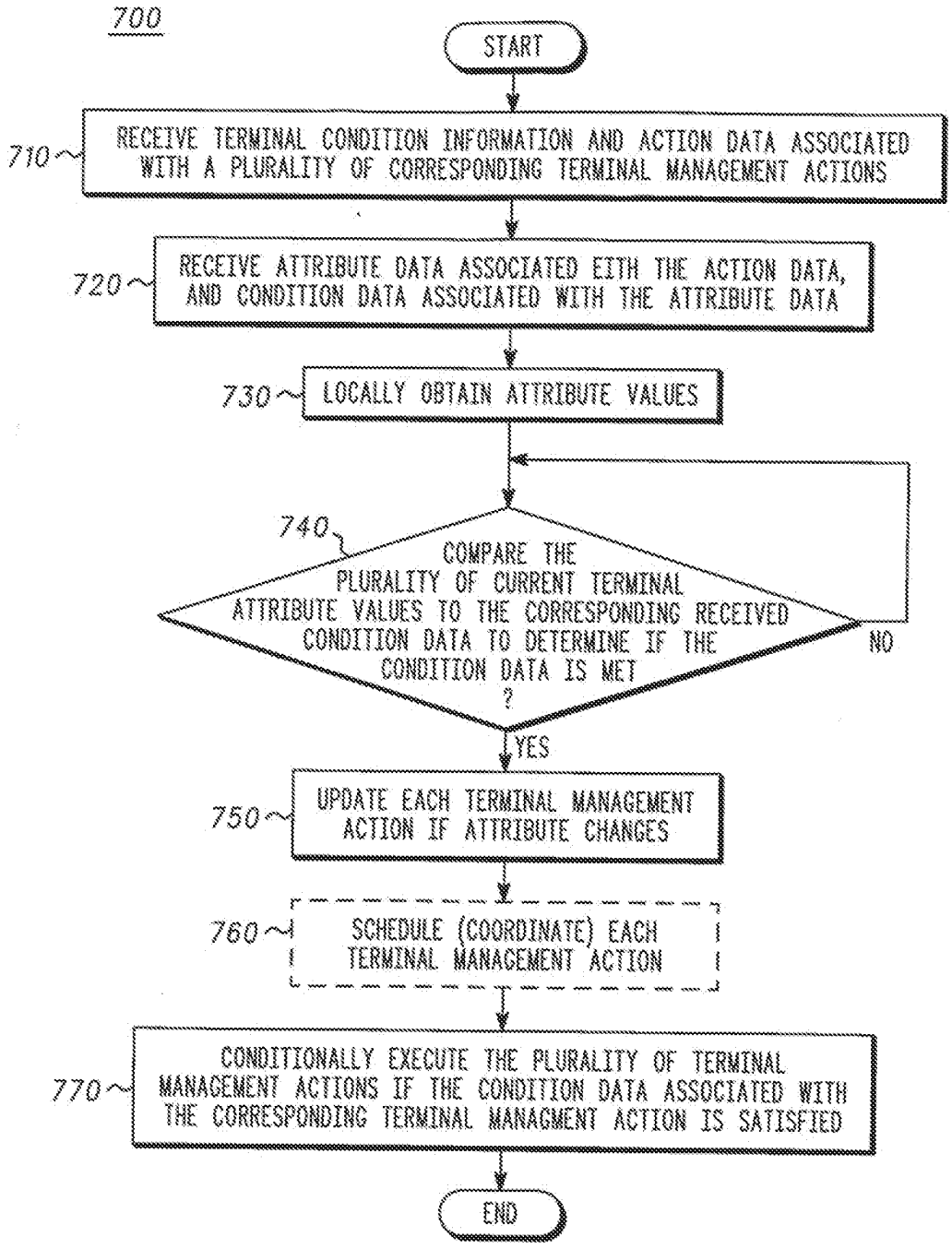

700
START
710
RECEIVE TERMINAL CONDITION INFORMATION AND ACTION DATA ASSOCIATED WITH A PLURALITY OF CORRESPONDING TERMINAL MANAGEMENT ACTIONS
720
RECEIVE ATTRIBUTE DATA ASSOCIATED EITH THE ACTION DATA, AND CONDITION DATA ASSOCIATED WITH THE ATTRIBUTE DATA
730
LOCALLY OBTAIN ATTRIBUTE VALUES
740
COMPARE THE PLURALITY OF CURRENT TERMINAL ATTRIBUTE VALUES TO THE CORRESPONDING RECEIVED CONDITION DATA TO DETERMINE IF THE CONDITION DATA IS MET ?
NO
YES
750
UPDATE EACH TERMINAL MANAGEMENT ACTION IF ATTRIBUTE CHANGES
760
SCHEDULE (COORDINATE) EACH TERMINAL MANAGEMENT ACTION
770
CONDITIONALLY EXECUTE THE PLURALITY OF TERMINAL MANAGEMENT ACTIONS IF THE CONDITION DATA ASSOCIATED WITH THE CORRESPONDING TERMINAL MANAGMENT ACTION IS SATISFIED
END

FIG. 7

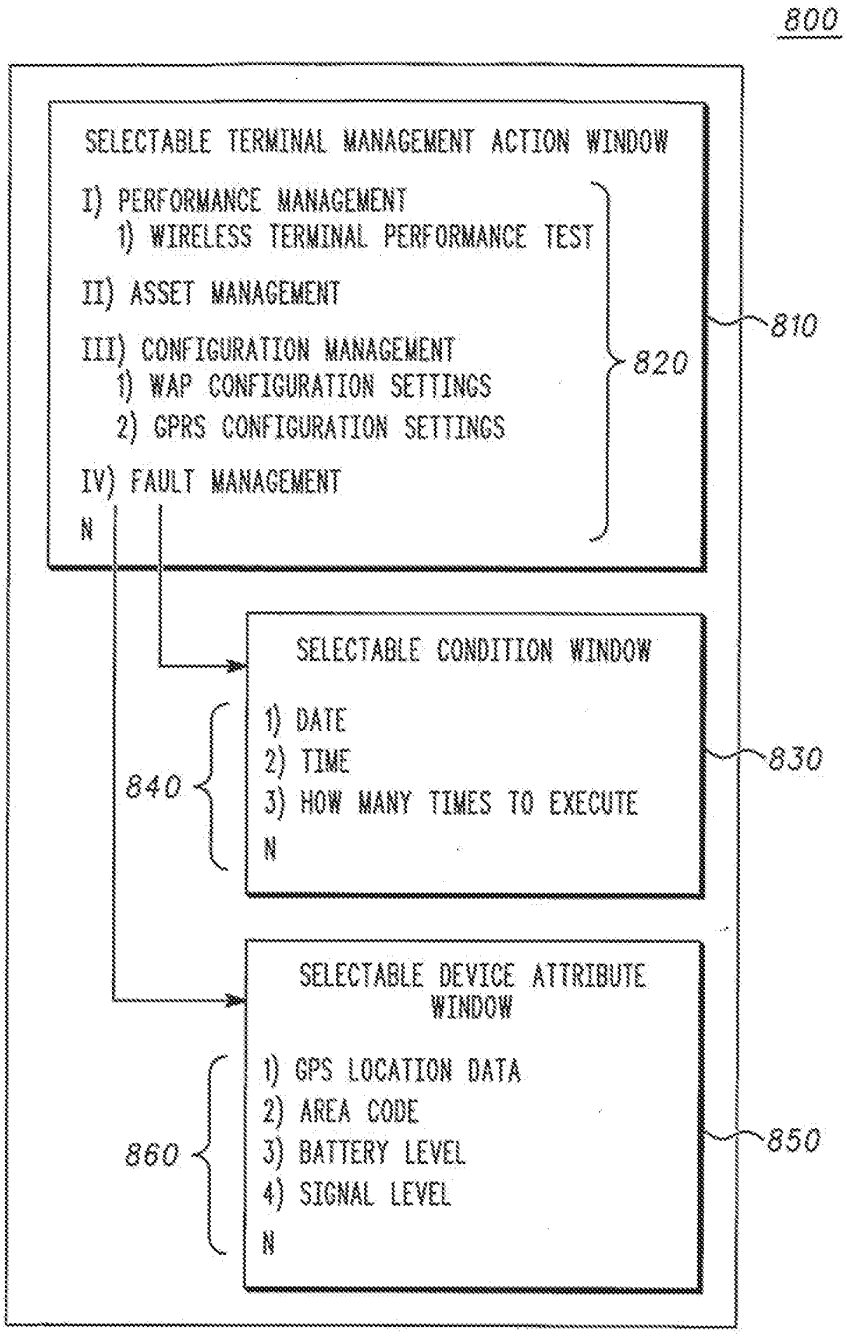

800
SELECTABLE TERMINAL MANAGEMENT ACTION WINDOW
I) PERFORMANCE MANAGEMENT
1) WIRELESS TERMINAL PERFORMANCE TEST
II) ASSET MANAGEMENT
III) CONFIGURATION MANAGEMENT
1) WAP CONFIGURATION SETTINGS
2) GPRS CONFIGURATION SETTINGS
IV) FAULT MANAGEMENT
N
820
810
SELECTABLE CONDITION WINDOW
1) DATE
2) TIME
3) HOW MANY TIMES TO EXECUTE
N
840
830
SELECTABLE DEVICE ATTRIBUTE WINDOW
1) GPS LOCATION DATA
2) AREA CODE
3) BATTERY LEVEL
4) SIGNAL LEVEL
N
860
850

FIG. 8

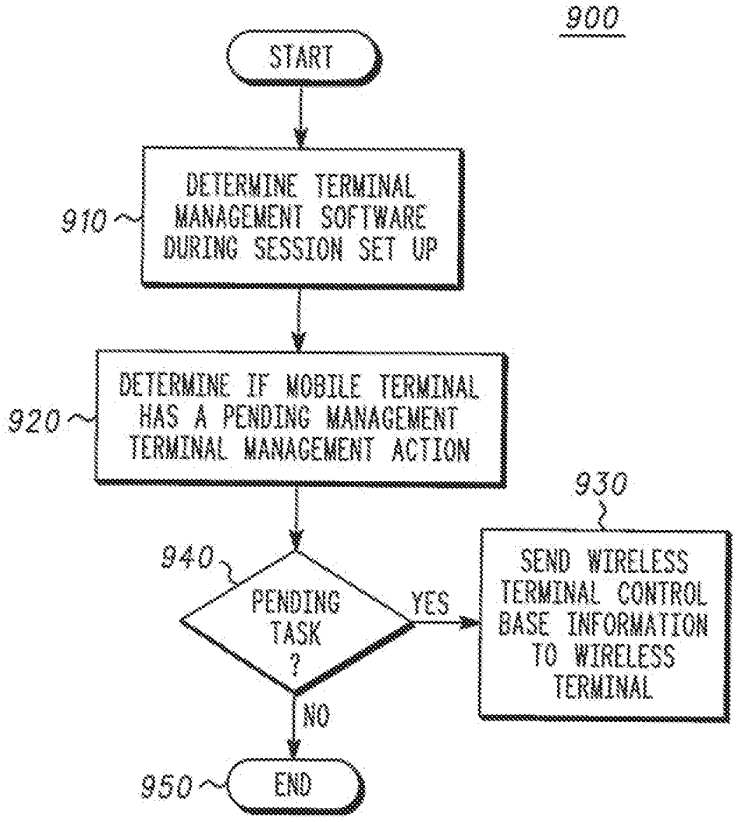

900
START
910
DETERMINE TERMINAL MANAGEMENT SOFTWARE DURING SESSION SET UP
920
DETERMINE IF MOBILE TERMINAL HAS A PENDING MANAGEMENT TERMINAL MANAGEMENT ACTION
930
940
PENDING TASK ?
YES
SEND WIRELESS TERMINAL CONTROL BASE INFORMATION TO WIRELESS TERMINAL
NO
950
END

FIG. 9

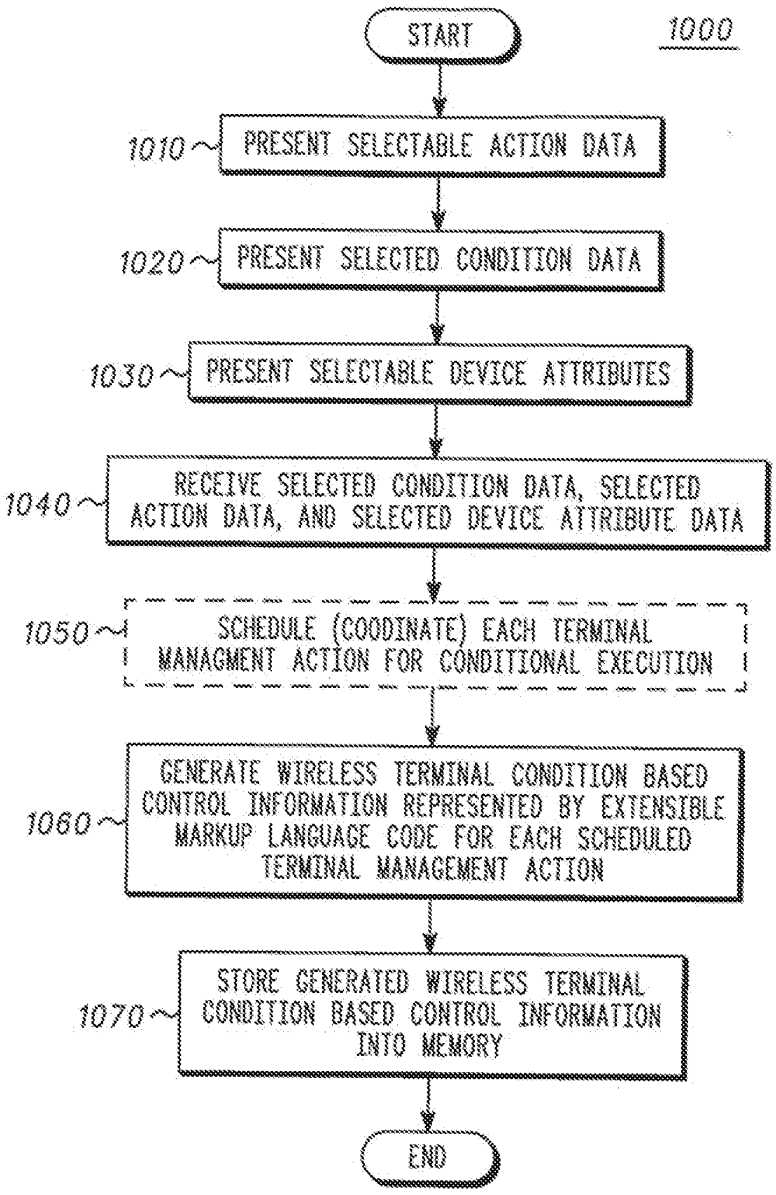

START
1000
1010
PRESENT SELECTABLE ACTION DATA
1020
PRESENT SELECTED CONDITION DATA
1030
PRESENT SELECTABLE DEVICE ATTRIBUTES
1040
RECEIVE SELECTED CONDITION DATA, SELECTED ACTION DATA, AND SELECTED DEVICE ATTRIBUTE DATA
1050
SCHEDULE (COODINATE) EACH TERMINAL MANAGMENT ACTION FOR CONDITIONAL EXECUTION
1060
GENERATE WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION REPRESENTED BY EXTENSIBLE MARKUP LANGUAGE CODE FOR EACH SCHEDULED TERMINAL MANAGEMENT ACTION
1070
STORE GENERATED WIRELESS TERMINAL CONDITION BASED CONTROL INFORMATION INTO MEMORY
END

FIG. 10

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number
WO 03/014891 A2

(51) International Patent Classification[7]: G06F 1/00

(21) International Application Number: PCT/US02/23226

(22) International Filing Date: 19 July 2002 (19.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/309,862 3 August 2001 (03.08.2001) US
09/992,310 19 November 2001 (19.11.2001) US

(71) Applicant: THE BOEING COMPANY [US/US]; P.O. Box 3707, M/S 13-08, Seattle, WA 98124-2207 (US).

(72) Inventor: ROCKWELL, Lawrence, L.; 8018-146th Avenue NE, Redmond, WA 98052 (US).

(74) Agent: GALBRAITH, Ann, K.; The Boeing Company, P.O. Box 3707, M/S 13-08, Seattle, WA 98124-2207 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/014891 A2

(54) Title: AN AIRBORNE SECURITY MANAGER

(57) Abstract: An airborne security management system is provided for monitoring security activities in a mobile network platform. The airborne security management system includes a mobile network that is interconnected via an unreliable communication link to a terrestrial-based network security management system; an intrusion detection system connected to the mobile network and operable to detect a security intrusion event whose origination is associated with the mobile network; and a mobile security manager adapted to receive the security intrusion events from the intrusion detection system. The mobile security manager is operable to transmit a message indicative of the security intrusion event to the network security management system and to perform security response activities in response to security commands received from the network security management system. The mobile security manager is further operable to command security response activities on the mobile network platform, when the mobile network platform is not connected with network security management system.

AN AIRBORNE SECURITY MANAGER

[0001] This application claims priority under 35 U.S.C. §119(e) to United States Provisional Application No. 60/309,862 filed on August 3, 2001, and entitled "An Airborne Security Manager" the specification and drawings of which are hereby expressly incorporated by reference.

TECHNICAL FIELD

[0002] The present invention relates generally to an airborne security management system for monitoring security activities in a mobile network platform, and more particularly to an autonomous airborne security manager for responding to detected security intrusion events when the mobile network platform is or is not in communication with a terrestrial-based network security management system.

BACKGROUND OF THE INVENTION

[0003] Broadband data and video services, on which our society and economy have grown to depend, have heretofore generally not been readily available to users onboard mobile network platforms such as aircraft, ships, trains, automobiles, etc. While the technology exists to deliver such services to most forms of mobile network platforms, past solutions have been generally quite expensive, with low data rates and/or available to only very limited markets of government/military users and some high-end maritime markets (i.e., cruise ships).

[0004] Previously developed systems which have attempted to provide data and video services to mobile network platforms have done so with only limited success. One major obstacle has been the high cost of access to such broadband data and video services. Another problem is the limited capacity of previously developed systems, which is insufficient for mobile network platforms carrying dozens, or even hundreds, of passengers who each may be simultaneously requesting different channels of programming or different data services. Furthermore, presently existing systems are generally not readily scalable to address the demands of the traveling public.

[0005] Of particular interest, presently existing systems also have not comprehensively addressed security issues relating to the mobile network platform. Therefore, it is desirable to provide a network security architecture for monitoring, reporting and responding to onboard security activities in a mobile network platform. It is envisioned that such a network security architecture should be designed to (a) secure computing resources to which passengers may have access on the mobile platform; (b) communicate reliably with terrestrial-based system components over an unreliable communication link; (c) provide a policy mediated response to detected security intrusion events occurring on the mobile platform; and (d) scale the management of the system to hundreds or thousands of mobile platforms.

SUMMARY OF THE INVENTION

[0006] In accordance with the present invention, an airborne security management system is provided for monitoring security activities in a mobile network platform. The security management system includes a mobile network that is interconnected via an unreliable communication link to a terrestrial-based network security management system; an intrusion detection system connected to the mobile network and operable to detect a security intrusion event whose origination is associated with the mobile network; and a mobile security manager adapted to receive the security intrusion events from the intrusion detection system. The mobile security manager is operable to transmit a message indicative of the security intrusion event to the network security management system and to perform security response activities in response to security commands received from the network security management system. The mobile security manager is further operable to command security response activities on the mobile network platform, when the mobile network platform is not connected with network security management system.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] The various advantages of the present invention will become apparent to one skilled in the art by reading the following specification and subjoined claims and by referencing the following drawings in which:

[0008] Figure 1 is a block diagram depicting a network security

architecture for a mobile network platform in accordance with the present invention;

[0009] Figures 2A and 2B are state machine diagrams illustrating a security policy for a given user access point on the mobile network platform in accordance with the present invention;

[0010] , Figure 3 is a diagram of an exemplary data structure for implementing the security policies of the present invention;

[0011] Figure 4 is a diagram depicting the primary software components of the network security architecture of the present invention;

[0012] Figure 5 is a block diagram depicting the functional software modules which comprise the airborne security manager in accordance with the present invention;

[0013] Figure 6 is a block diagram depicting the functional components implementing the terrestrial control and data storage functions of a terrestrial-based network security system in accordance with the present invention;

[0014] Figure 7 is an exemplary aircraft browser window used to implement the monitoring and manual control functions of a terrestrial-based network security system in accordance with the present invention; and

[0015] Figure 8 is an exemplary aircraft status window used to implement the monitoring and manual control functions of a terrestrial-based network security system in accordance with the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0016] Figure 1 illustrates a network security architecture 10 for

monitoring security activities in an unattended mobile network platform 12. The primary purpose of the network security architecture 10 is to monitor, record, report and respond to security-relevant events associated with the mobile network platform 12. In a preferred embodiment, the network security architecture 10 supports a mobile network platform residing in an aircraft. The mobile network platform 12 is in turn interconnected via one or more unreliable wireless communication links 14 to a terrestrial-based communication system16, including a terrestrial-based network security management system 18. While the following description is provide with reference to an airborne application, it is readily understood that the broad aspects of the network security architecture are applicable to mobile network platforms which may reside in passenger buses, cruise ships, etc.

[0017] It is envisioned that the mobile network platform 12 provides aircraft passengers a suite of broadband two-way data and video communication services. The infrastructure allows information to be transferred to and from the aircraft at high enough data rates to support a variety of services. To do so, the mobile network platform 12 is primarily comprised of four subsystems: an antenna subsystem 22, a receive and transmit subsystem (RTS) 24, a control subsystem 26, and a cabin distribution subsystem 28. Each of these four subsystems will be further described below.

[0018] The antenna subsystem 22 provides two-way broadband data connectivity and direct broadcast television reception capability to the aircraft. Although the invention is not limited thereto, the antenna subsystem 22 is generally designed to provide this connectivity during cruise conditions (limited roll

and pitch angles) of the aircraft. Connectivity with the aircraft is most commonly achieved via a K band Fixed Satellite Service (FSS) satellite, a Broadcast Satellite Service (BSS) satellites, and/or a direct broadcast television service (DBS) satellite.

[0019] For illustration purposes, additional description is provided for the processing associated with Ku band satellite broadcast signals. The antenna subsystem 22 may receive and/or transmit Ku band satellite broadcast signals. The antenna system 22 down-converts an incoming Ku-band signal, amplifies, and outputs the L-band signals to the RTS 24. The antenna system may also provide a broadband downlink capability. In this case, the antenna system 22 receives an L-band data signal from an on-aircraft modem, up-converts this signal, amplifies it and then broadcasts as a Ku band signal to selected satellite transponders.

[0020] The receive and transmit subsystem (RTS) 24 operates in receive and transmit modes. In receive mode, the RTS 24 may receive rebroadcast video signals, rebroadcast audio signals and/or IP data embedded in an L-band carrier. The RTS 24 in turn demodulates, de-spreads, decodes, and routes the received signals to the cabin distribution subsystem 28. In transmit mode, the RTS 24 sends IP data modulated into an L-band signal. The RTS 24 encodes, spreads, and modulates the signal the IP data it receives from the cabin distribution subsystem 28.

[0021] The control subsystem 26 controls the operation of the mobile security platform 12 and each of its four subsystems. Of particular interest, the control subsystem 26 includes one or more intrusion detection subsystems 32 and

an airborne security manager 34. An intrusion detection subsystem 32 is operable to detect security intrusion activities which may occur on or in relation to the mobile network platform. To do so, an intrusion detection subsystem 32 inspects all of the data packets entering a computing device on which it is hosted and, upon detection of a security intrusion activity, transmits a security intrusion event to the airborne security manager 34. As will be apparent to one skilled in the art, the intrusion detection subsystem 32 may be implemented using one of many commercially available software products.

[0022] The airborne security manager 34 is responsible for enforcing security policy for an aircraft. Because communication with the aircraft may be sporadic, the airborne security manager 34 must provide the capability to act autonomously when responding to security intrusion events. When a security intrusion event is detected, the airborne security manager 34 responds appropriately in accordance with a customizable security policy. Thus, the airborne security manager 34 is adapted to receive security intrusion events from any of the intrusion detection subsystems and operable to implement a security response. Exemplary responses may include warnings one or more passengers on the aircraft, alerting terrestrial-based security administrators, and/or disconnecting a passenger's network access.

[0023] The cabin distribution subsystem (CDS) 28 provides network connectivity through a plurality of user access points to the passengers of the aircraft. In a preferred embodiment, the cabin distribution system may be composed of either a series of 802.3 Ethernet switches or 802.11X wireless access points. It should be noted that the current 802.11B standard only allows

for a shared secret between all users of a wireless access point and thus is not suitable for providing the desired level of communication privacy in the passenger cabin. In contrast, next generation wireless standards, such as 802.11X ("X" denotes a revision of 802.11 beyond "B") will support "channelized" or individual user level encryption. It is envisioned that such wireless standards are within the scope of the present invention.

[0024] Each user access point preferably has the properties of a managed layer 3 switch. First, each user access point must enforce the association of IP address and MAC address with a particular port. This requirement his applicable to either a wired and wireless cabin environment. A second requirement for each user access point is to accept a command to shut off its access port. In the case of a wireless access device, a communication channel consisting of a particular frequency, time division or sub-frame substitutes for the physical access port. A third requirement for each user access point is to preclude passengers from eavesdropping or receiving Ethernet packets not directly addressed to them. In a wired cabin distribution system, this can be accomplished through the use of a switched Ethernet architecture. In a wireless cabin distribution system, this can be accomplished through the use of "channel level encryption" specific to a particular user.

[0025] The design of a security policy mechanism is the most fundamental element of the network security architecture 10. In accordance with the present invention, it is envisioned that the security policy will be designed within the following design constraints. First, the security policy mechanism should map different security intrusion events to different responses. It should be

appreciated that the severity of response is based on the danger of the detected activities. Second, the automated response policy has to be enforced at all times (subject to over-ride conditions), regardless of whether airborne to terrestrial communications are available or not. If the automated responses are disabled during periods of connectivity, the connectivity might fail before a security administrator has a chance to take action in which case the system reverts to the automated policy in effect prior to the override. The security administrator can retract the response if they desire. Third, the policy mechanism has to arbitrate between automated responses from the airborne security manager and manual commands received from terrestrial-based security administrators. If the automated system mistakenly blocks a passenger's network address, and the terrestrial administrator overrides that action, the security policy mechanism needs to know about that action and not try to enforce the block.

[0026] State machines are a flexible, yet intuitively appealing, mechanism for modeling complex behaviors. Therefore, state-machines have been chosen to represent the security policies of the present invention. Figures 2A and 2B illustrates basic UML state machines which model the security policy associated with an user access point in the mobile network platform.

[0027] In Figure 2A, each user access point can be in one of three defined states. By default, all user access points begin in a normal state 42. A security intrusion event of any kind will result in a transition to either a suspected state 44 or a disconnected state 46 for the applicable user access point. Each transition is in the form of "event/response" where events are the external triggers that cause the state transition and responses are external actions that the system

initiates when making the transition. For instance, a low or medium priority event 48 occurring in a normal state will cause the system to log the event and/or attempt to provide a warning to the passenger connected at that user access point. The user access point then transitions to the suspected state as shown in Figure 2A.

State machine models may be enhanced to incorporate manual controls. Specific manual control commands enable a terrestrial-based security administrator to explicitly disable or enable a user access point from the ground. By adding a state that indicates that the user access point is under manual control ensures that the automated responses do not override the manual control command received from the security administrator. Therefore, it is envisioned that each state machine may provide an autoresponse disable state 50 as shown in Figure 2B. Transitions to and from the autoresponse disable state are commanded by a terrestrially-based security administrator. While in the autoresponse disable state, the administrator can initiate any one of various predefined security responses. In the event connectivity is lost between the administrator and the aircraft, the state machine model reverts to the normal state or the previous state depending on configuration settings.

State machines models are also used to represent each of the host servers or other types of computing devices which reside on the mobile security platform. In this way, a server that is under attack may respond differently than a user access point. It is also envisioned that each of the state machines can be tied together through synthetic event generation, such that when a server is under attack, the user access points may employ a different security policy that is less

tolerant of suspicious behavior.

[0028] Each state machine can be represented by a data structure 51 as depicted in Figure 3. The data structure includes a current state 52, a possible security event 54, a resulting state 56 and a possible response 58. In this way, each state can be cross-referenced against possible events to produce a resulting state and a list of possible actions. Possible events may include (but are not limited to) a security intrusion event having high priority, a security intrusion event having medium priority, a security intrusion event having a low priority, a reset event, a timer expiration event, a communication link up event, a communication link down event and one or more custom events for supporting manual control commands from the security administrator. Possible responses may include (but are not limited to) setting a timer, installing a filter, resetting a filter, alerting control panel, alerting terrestrial-based security administrator, disconnecting user access point, issuing a passenger warning, and one or more predefined customer responses. One skilled in the art will readily recognize from such discussion how to implement a security policy mechanism in accordance with the present invention.

[0029] Referring to Figure 4, the overall network security architecture 10 may be logically decomposed into five major components. The five major components are airborne policy enforcement 62, air-ground communication 64, terrestrial control and data storage 66, terrestrial monitoring and manual control 68, and terrestrial policy editing and assignment 70. Each of these logical components are also mapped to their physical location within the network security architecture 10 as shown in Figure 4.

[0030] The airborne policy enforcement component 62 is provided by the airborne security manager 34. The primary responsibilities of the airborne security manager include (but are not limited to) managing and monitoring intrusion detection sensors, monitoring other airborne event sources, responding to security events in accordance with the applicable security policy, monitoring the airborne intrusion detection sensors, configuring static network traffic filters at user access points, executing any manual overrides commands from the terrestrial-based network security management system, installing new security policies received from the terrestrial-based network security management system, and reporting events and status of interest to the terrestrial-based network security management system. As will be apparent to one skilled in the art, the airborne security manager 34 is comprised of one or more software applications residing on one or more server(s) on each aircraft. A configuration of redundant airborne security managers provide for fail over in the event of a hardware or software failure.

[0031] With reference to Figure 5, the airborne security manager 34 is further comprised of five functional modules: an event response module 72, an onboard status module 74, a policy manager 76, a persistent storage manager 78, and a communication manager 80. The event response module 72 is responsible for receiving events, interpreting the active security policy, and triggering the appropriate actions in response to each event. It should be appreciated that this module is adapted to handle events other than security intrusion events received from the intrusion detection subsystems.

[0032] In conjunction with the onboard status module 74, the event

response module interprets and executes the state machine representing the active security policy. For instance, upon arrival of a security intrusion event, the event response module determines whether the event is associated with an individual passenger connection, an individual host server, or the airborne security manager as a whole. This module then retrieves the current state of that passenger connection, host server, or airborne security manager from the onboard status module 74 and performs the actions associated with that state and event in accordance with the active security policy. Exemplary actions may include issuing new events, making state transitions, modifying network filters, disabling passenger connections, and/or queuing messages for transmission to the terrestrial-based network security management system.

[0033] The onboard status module 74 maintains the current state of each individual passenger connection, each host server, and of the airborne security manager as a whole for the purpose of directing the state machine event response. The onboard status module 74 also tracks the status of intrusion detection sensors (e.g., signature file, operational/inactive status, sensor configuration) as well as collects status information from the other onboard modules.

[0034] The policy manager 76 is responsible for reacting to commands from the terrestrial-based network security system regarding security policy loading and activation. The policy manager also serves as a repository for configuration information relating to the airborne security manager, including, for instance, general communications parameters that determine frequency of status reports and event reporting.

[0035] The persistent storage manager 78 manages the overall data storage requirements for the onboard network security architecture. Data residing in persistent storage generally falls into one of three categories: (1) communications queue (i.e., messages to be transmitted to the terrestrial-based security management system), (2) onboard status (i.e., per-passenger connection, per-host, and system-wide data requirements), and (3) security policies. The persistent storage manager may rely on various well known, lightweight mechanisms for data storage.

[0036] Referring to Figure 4, the terrestrial control and data storage (C&DS) component 66 is provided by the terrestrial-based network security management system 16. The control and data storage functions include (but are not limited to) storing all event data in persistent storage, tracking the desired and last known configurations for each aircraft, supporting multiple security management consoles having multiple windows, notifying open console windows of any data changes that affect the window contents, providing an interface for effecting manual overrides in security policy, offering a reporting interface for reviewing stored data, and controlling access to all stored data. This component may be implemented using Java-based applications residing on one or more terrestrial servers which constitute the network security management system 16.

[0037] A more detailed description of the terrestrial control and data storage component 66 is provided with reference to Figure 6. This terrestrial component will maintain one aircraft object 90 for each aircraft associated with the security architecture. The aircraft object 90 maintains all state information for a given aircraft as well as keeps track of the last reported and the desired state of

the airborne security manager 34 residing on the given aircraft. The aircraft object 90 is a dynamic object, such that it state is maintained in dynamic memory and can be reconstructed from event histories, if necessary. Any activity that could alter the state of the airborne security manager 34 is performed by invoking a method of the aircraft object. Each method represents an event and is logged in an appropriate event log. In addition, these methods are all synchronized, ensuring that only one thread can be effecting state changes at any given time. In order to eliminate the possibility of deadlock, none of these event operations will block on communication or issue events to other aircraft.

[0038] The aircraft object 90 uses a communication subsystem 100 to exchange information with the airborne security manager 34. The aircraft object 90 issues commands and requests for status reports as well as receives events and status reports. Until an appropriate event or status report is received, any command is considered pending. This does not mean that the command has not yet executed – it may not have been executed, or it may have been and the acknowledging status report has simply not yet been received. Due to this gap in knowledge about what is actually taking place onboard the aircraft, the aircraft object 90 must carefully differentiate between the last known status and the desired status.

[0039] The aircraft object 90 is the controller in a Model-View-Controller architecture as is well known in the art. In this paradigm, the model is the data stored in a database, and the views are the various user interfaces being used to display information about the aircraft. The aircraft object is responsible for updating all of the views any time the model changes. In order to enforce this, all

changes to the model must be performed by the aircraft object and the aircraft object must keep track of those user interfaces that could be affected by the change.

[0040] The aircraft object 90 also maintains a collection of host objects 92 and passenger connection objects 94. The host objects 92 are used to represent the state of each onboard host server that the airborne security manager 34 is responsible for. The passenger connection objects 94 represent the individual passenger connections to the onboard network.

[0041] The terrestrial control and data storage component 66 also includes a single aircraft container object 96. It is envisioned that this object may be implemented as a collection class, such as a hash table. Under this approach, aircraft objects will be created by the aircraft container 96 for every aircraft in the system. By routing incoming communication through the aircraft container 96, we ensure that the communication subsystem 100 will be able to deliver incoming messages to the appropriate aircraft object. In addition, the container concept may be used to facilitate the manner in which aircraft objects are created. For instance, aircraft objects may be created only as they are needed. When an incoming message is received, the aircraft container 96 locates the applicable aircraft object. If the aircraft object is not present in memory, the aircraft container can create the object. Likewise, aircraft objects that are no longer being actively monitored could be deleted until they are needed again.

[0042] The terrestrial control and data storage component 66 will also maintain event histories for each of the aircraft in a central database 98. The database 98 will maintain a record of all the events reported by an aircraft in the

system. In addition, it will maintain a record of all of the commands performed by terrestrial-based security administrator. The former represents the last known state of each aircraft; whereas the latter represents the desired stated of each aircraft. The choice of the term "last known" reflects the time delay between events occurring on board the aircraft which might not have been reflected on the ground.

[0043] Security policy files are also stored within the database 98. As a configuration option, in order to maintain a history of old policies, the policy tables may be append-only. The primary policy table will maintain a mapping of names and version numbers to a series of smaller policy elements. The communication subsystem 100 interfaces with the database 98 in order to retrieve security policy files and update the policy files onboard the aircraft.

[0044] A policy manager 99 will be responsible for any changes to policy files. This object is necessary because policy is the only thing that is not associated with a single aircraft. The policy manager 99 will ensure that any changes to policy files are properly versioned. It will also be responsible for delivering updated policy to one or more aircraft.

[0045] Returning to Figure 4, the terrestrial monitoring and manual control component 68 and the terrestrial policy editing and assignment component 70 also reside at the terrestrial-based network security management system 12. The monitoring and manual control component functions include (but are not limited to) monitoring the state and activities of a group of aircraft and selecting an individual aircraft for closing examination, monitoring the state and activities of a single aircraft and selecting an individual server or passenger connection for

closer examination, monitoring the state and activities of a single airborne server, manually controlling a single airborne server, monitoring the state and activities of a single airborne passenger connection, and manually controlling a single airborne passenger connection. This component may be implemented using a Java-based user interface running on one or more terrestrial servers.

[0046] To support the monitoring and manual control functions, the user interface includes a number of windows that may be monitored by a human network security administrator. For instance, an aircraft browser allows groups of aircraft to be navigated and aggregate/summary information displayed as shown in Figure 7. However, this window does not show the status of the communication link. In order to display such status information, the user can select a specific aircraft from the aircraft browser, thereby navigating to an aircraft status window. An exemplary aircraft status window is shown in Figure 8. The aircraft status window enables the user to view all data relevant to a specific aircraft in a single tree structure view 102. In addition, all logged events and commands are displayed in a lower log panel 104. The tabs 106 along the top of the window permit navigation to other panels which in turn focus on a different specific element associated with the aircraft. For instance, the seat panel 108 will provide status information, log detail, and manual controls for a specific seat. Other exemplary windows used to support the monitoring and manual control functions may include (but is not limited to) a passenger connection status window that focuses on displaying information for a single passenger connection, an onboard host status window that focuses on displaying information on a specific host computing device residing on the aircraft, and an events log window that displays

event information for a given group, aircraft, passenger connection or host device. It is envisioned that the above-described windows are merely representative of some of the functionality and appearance that be used to implement the monitoring and manual control functions of the present invention.

[0047] In addition to monitoring and manual control, services for editing security policy files and distributing security policy updates also reside at the terrestrial-based network security management system 16. The policy editing and application functions include (but are not limited to) editing sensor configuration files, retrieving intrusion detection signature file updates from the applicable vendor website, editing response policy state machines and parameters, editing static security configurations, combining sensor files, signature files, response policies, and static configuration into specific security policies, providing version control over security policy updates, browsing the aircraft in the system by last known policy and desired policy, and distributing a new policy to a selected group of aircraft. The editing of security policy is not intended to be a routine daily activity. For this reason, policy editing and application functions are treated as a separate, distinct logical component from the other functions administered through the user interface running on the terrestrial servers.

[0048] The air-ground communication component 64 is responsible for communication between the airborne security manager and the terrestrial servers. Thus, this component is distributed across these two physical locations. The air-ground communication functions include (but are not limited to) providing non-blocking communications, retrying transmissions until reliable delivery is achieved, queuing up messages during periods of non-connectivity, handling communication

session authentication, utilizing cryptographic integrity checks to protect against tampering and replay, optimizing away redundant or superseded messages where possible, utilizing available bandwidth according to message priorities, minimizing bandwidth consumption, and delivering security policy updates to aircrafts. Logically isolating the communications component helps protect the design of the airborne security manager and the terrestrial servers from unnecessary complexity arising from sporadic connectivity.

[0049] The foregoing discussion discloses and describes preferred embodiments of the invention. One skilled in the art will readily recognize from such discussion, and from the accompanying drawings and claims, that changes and modifications can be made to the invention without departing from the true spirit and fair scope of the invention as defined in the following claims.

What is claimed is:

1. A network security architecture for monitoring security activities in a mobile network platform, comprising

a mobile network residing on the mobile network platform, the mobile network being interconnected via an unreliable communication link to a terrestrial-based network security management system;

an intrusion detection system connected to the mobile network and residing on the mobile network platform, the intrusion detection system operable to detect a security intrusion event that is associated with the mobile network; and

a mobile security manager residing on the mobile network platform and adapted to receive the security intrusion events from the intrusion detection system, the mobile security manager is further operable to perform security response activities in response to the security intrusion events, when the mobile network platform is not connected with network security management system.

2. The network security architecture of Claim 1 wherein the mobile security manager is operable perform security response activities in accordance with a security policy resident on the mobile network platform.

3. The network security architecture of Claim 2 wherein the security policy is defined as a plurality of predefined security intrusion events and a

corresponding security response for each of said plurality of security intrusion events.

4. The network security architecture of Claim 2 wherein the security policy is defined by a data structure having a current operational state element, a possible security intrusion event element, a resulting operational state element, and a security response element.

5. The network security architecture of Claim 1 wherein the mobile network includes a plurality of user access points, such that the security intrusion event is associated with one of the plurality of user access points and the security response is directed to said one of the plurality of user access points.

6. The network security architecture of Claim 5 wherein the security response is selected from the group consisting of logging the security intrusion event received from the intrusion detection system, providing a warning message to at least one of said user access points, providing an alert message to a terrestrial-based network security management system, installing a network traffic blocking filter at one of said user access points, and disconnecting one of said user access points from the mobile network.

7. The network security architecture of Claim 5 wherein the mobile security manager maintains an indicator of the current operational state for each

of the plurality of user access points, such that the security response directed to said one of the plurality of user access points is in part based on the operational state of said one of the plurality of user access points.

8. The network security architecture of Claim 7 wherein the current operational state for any given user access point is selected from the group consisting of a normal state, a suspected state, and a disconnect state.

9. The network security architecture of Claim 7 wherein the mobile security manager is further operable to identify the current operational state for said one of the plurality of user access points and perform security response activities based in part on the identified operational state and the security intrusion event received from the intrusion detection system.

10. The network security architecture of Claim 9 wherein the mobile security manager is further operable to modify the current operational state for said one of the plurality of user access points in accordance with the security policy.

11. The network security architecture of Claim 1 wherein the mobile security manager is operable to transmit a message indicative of the security intrusion event to the network security management system and to perform security response activities in response to security commands received from the

network security management system.

12. A method for monitoring security activities associated with a network residing in a mobile network platform, the mobile network platform being interconnected via an unreliable communication link to a terrestrial-based network security management system, comprising:

detecting a security intrusion event whose origination is associated with the network residing on the mobile network platform;

providing a mobile security manager residing on the mobile network platform, where the mobile security manager is adapted to receive the security intrusion event; and

performing a security response activity in response to the detected security intrusion event, when the mobile network platform is not connected with the network security management system.

13. The method of Claim 12 wherein the step of performing a security response activity further comprises applying the security response activity in accordance with a security policy, where the security policy is defined as a plurality of predefined security intrusion events and a corresponding security response for each of said plurality of security intrusion events.

14. The method of Claim 12 further comprising the steps of applying the security response activity in accordance with a security policy, where the security policy is defined by a data structure having a current operational state element, a possible security intrusion event element, a resulting operational state element, and a security response element.

15. The method of Claim 12 wherein the network includes a plurality of user access points, such that the security intrusion event is associated with one of the plurality of user access points and the security response is directed to said one of the plurality of user access points.

16. The method of Claim 15 wherein the security response activity is selected from the group consisting of logging the security intrusion event, providing a warning message to at least one of the user access points, providing an alert message to a terrestrial-based network security management system, installing a network traffic blocking filter at one of the user access points, and disconnecting one of the user access points from the network.

17. The method of Claim 16 further comprising the steps of maintaining an indicator of the current operational state for each of the plurality of user access points and performing a security response activity in response to the detected security intrusion event, where the security response activity is in part based on the operational state of said one of the plurality of user access points.

18. The method of Claim 17 wherein the current operational state for any given user access point is selected from the group consisting of a normal state, a suspected state, and a disconnect state.

19. An airborne security system for monitoring security activities associated with a network residing on an aircraft, the aircraft being interconnected via an unreliable communication link to a terrestrial-based network security management system, comprising:

an intrusion detection system connected to the network and operable to detect a security intrusion event that is associated with network; and

an airborne security manager connected to the network and adapted to receive the security intrusion event from the intrusion detection system, the security manager is further operable to perform security response activities in accordance with a security policy, when the aircraft is not connected with the network security management system.

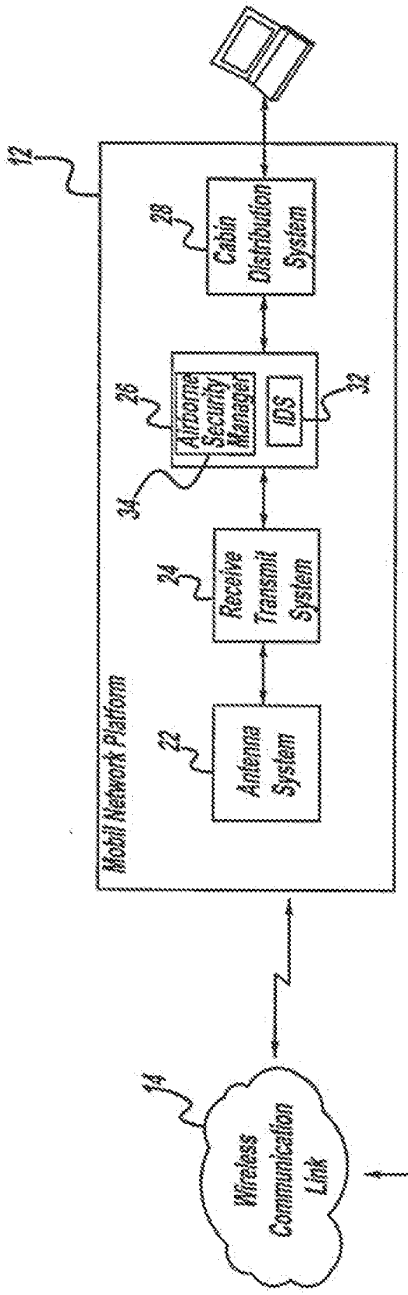

12
Mobil Network Platform
22
Antenna System
24
Receive Transmit System
26
Antenna Security Manager
32
IDS
28
Cabin Distribution System
34
14
Wireless Communication Link
18
Terrestrial Network Security Management System
Ground Comm System
16

Figure - 1

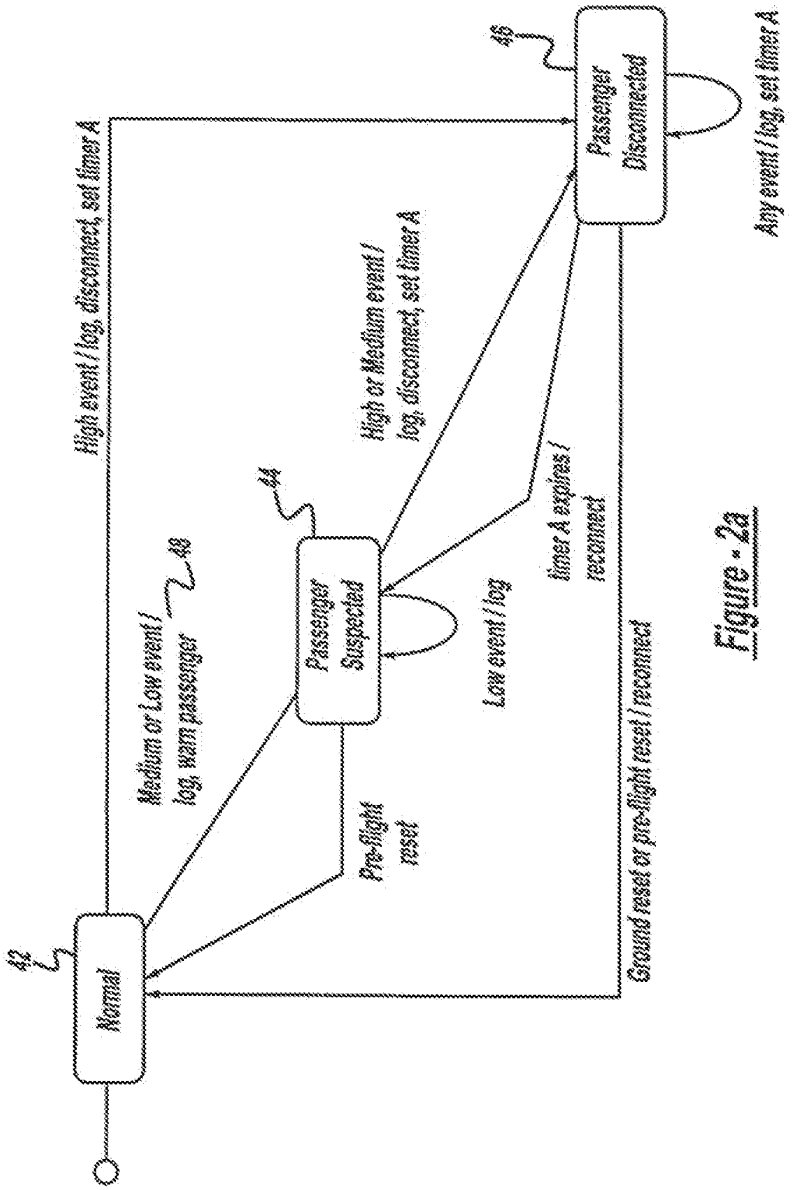

42
Normal
44
Passenger Suspected
46
Passenger Disconnected
High event / log, disconnect, set timer A
Medium or Low event / log, warn passenger
48
Pre-flight reset
Low event / log
High or Medium event / log, disconnect, set timer A
timer A expires / reconnect
Any event / log, set timer A
Ground reset or pre-flight reset / reconnect

Figure - 2a

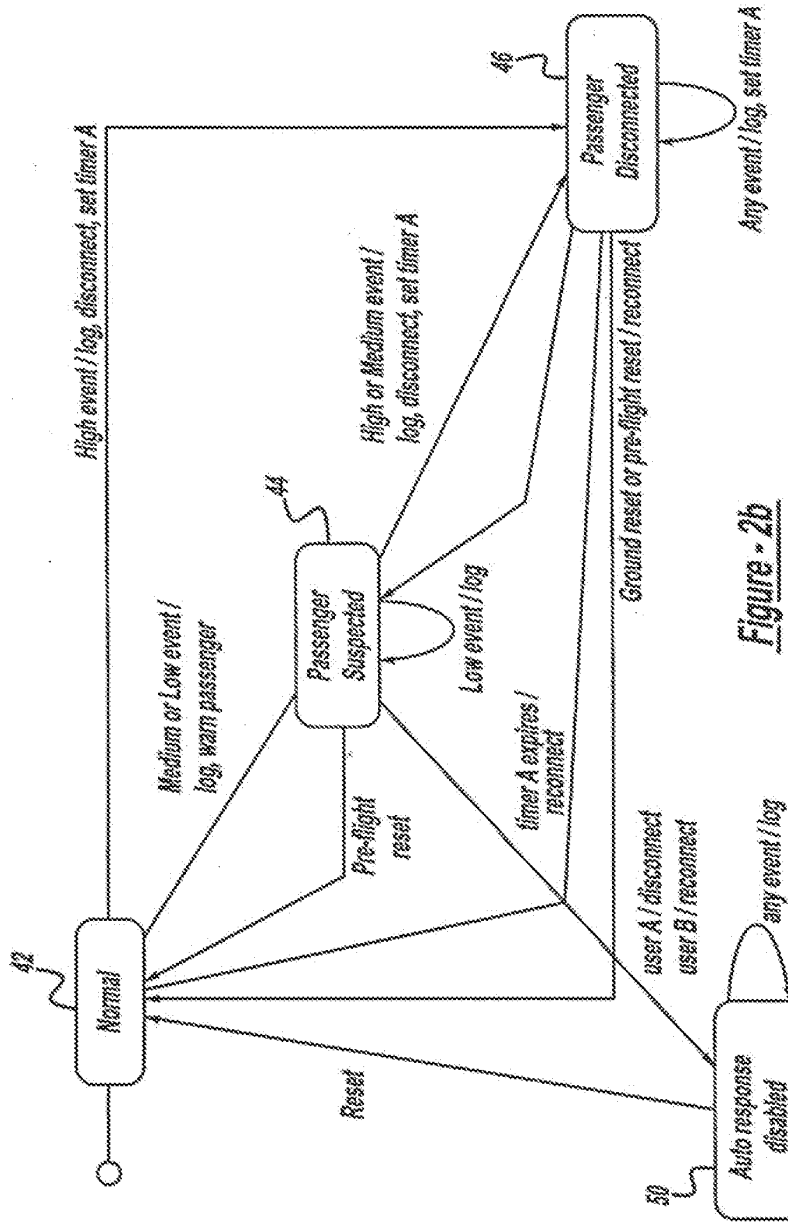

42
Normal
44
Passenger Suspected
46
Passenger Disconnected
50
Auto response disabled
High event / log, disconnect, set timer A
Medium or Low event / log, warn passenger
High or Medium event / log, disconnect, set timer A
Low event / log
Pre-flight reset
timer A expires / reconnect
Ground reset or pre-flight reset / reconnect
Any event / log, set timer A
user A / disconnect user B / reconnect
any event / log
Reset

Figure - 2b

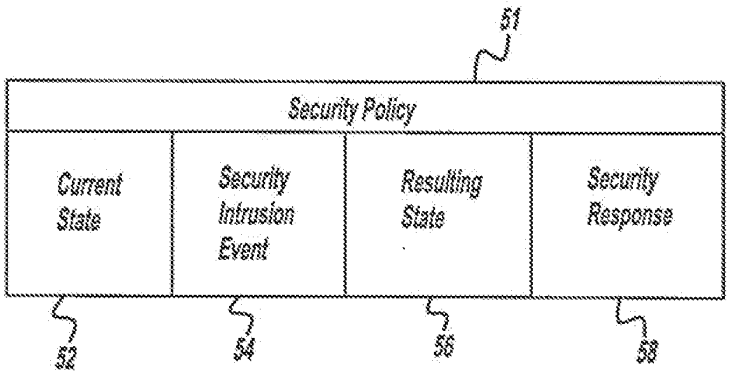

51
Security Policy
Current State
Security Intrusion Event
Resulting State
Security Response
52
54
56
58

Figure - 3

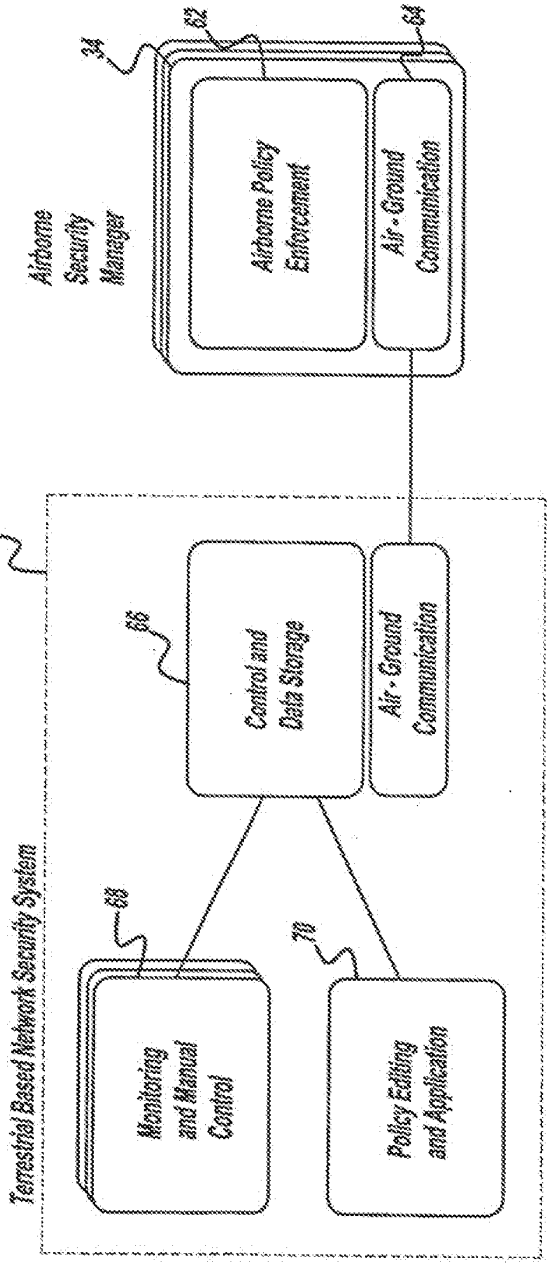

Airborne Security Manager
34
62
64
Airborne Policy Enforcement
Air - Ground Communication
16
Terrestrial Based Network Security System
66
Control and Data Storage
Air - Ground Communication
68
Monitoring and Manual Control
70
Policy Editing and Application

Figure - 4

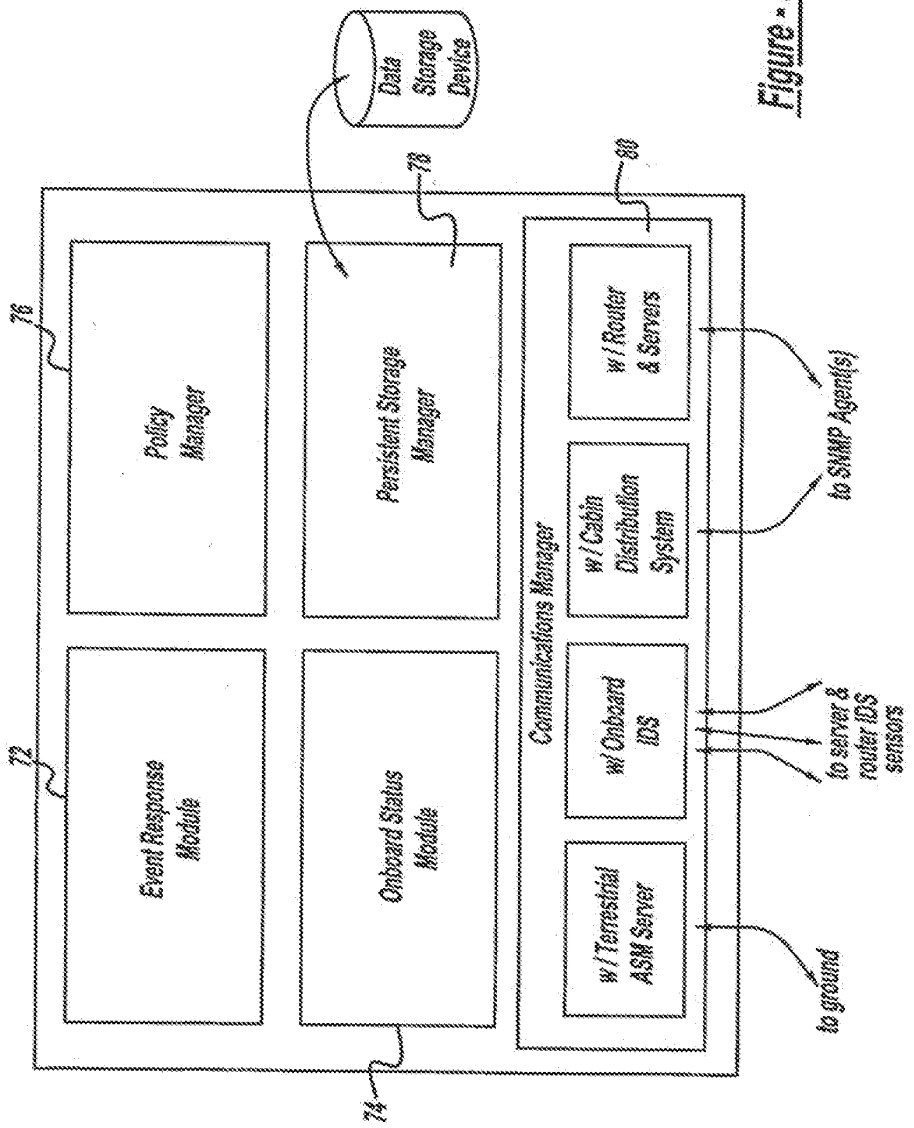

Data Storage Device
78
80
76
Policy Manager
Persistent Storage Manager
w/ Router & Servers
w/ Cabin Distribution System
to SNMP Agent(s)
Communications Manager
w/ Onboard IDS
to server & router IDS sensors
72
Event Response Module
Onboard Status Module
w/ Terrestrial ASM Server
to ground
74

Figure - 5

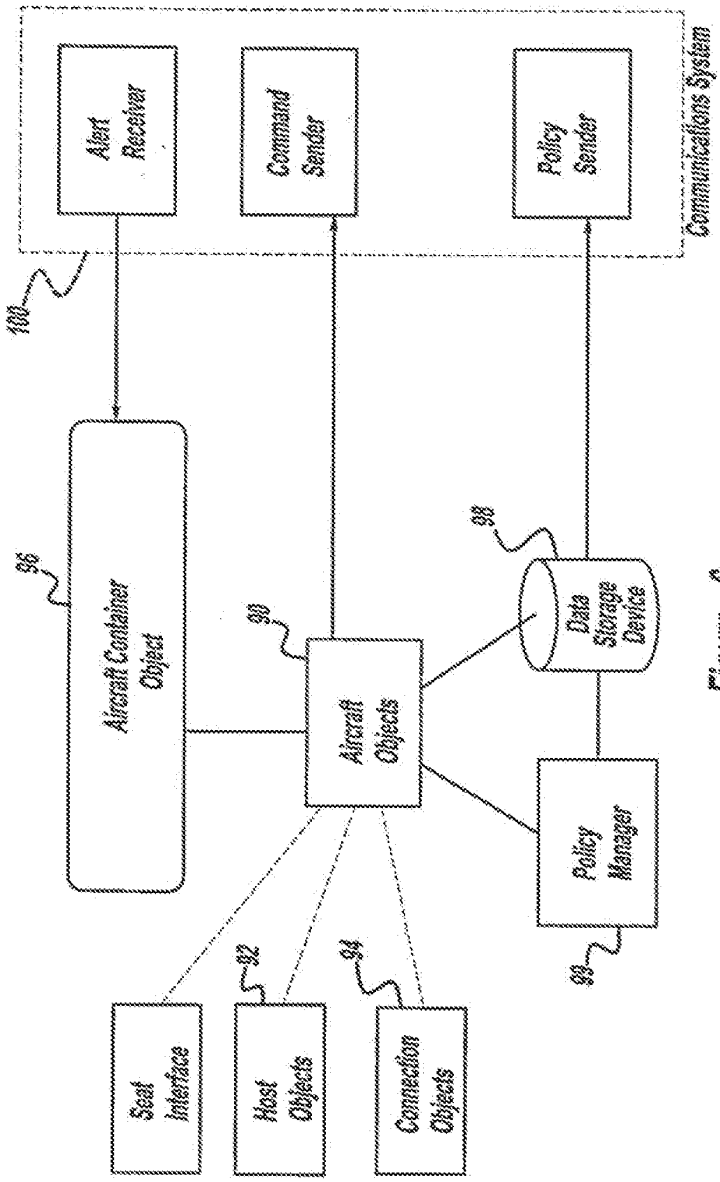

Communications System
Alert Receiver
Command Sender
Policy Sender
100
Aircraft Container Object
96
Aircraft Objects
90
Data Storage Device
98
Policy Manager
93
Seat Interface
Host Objects
92
Connection Objects
94

Figure - 6

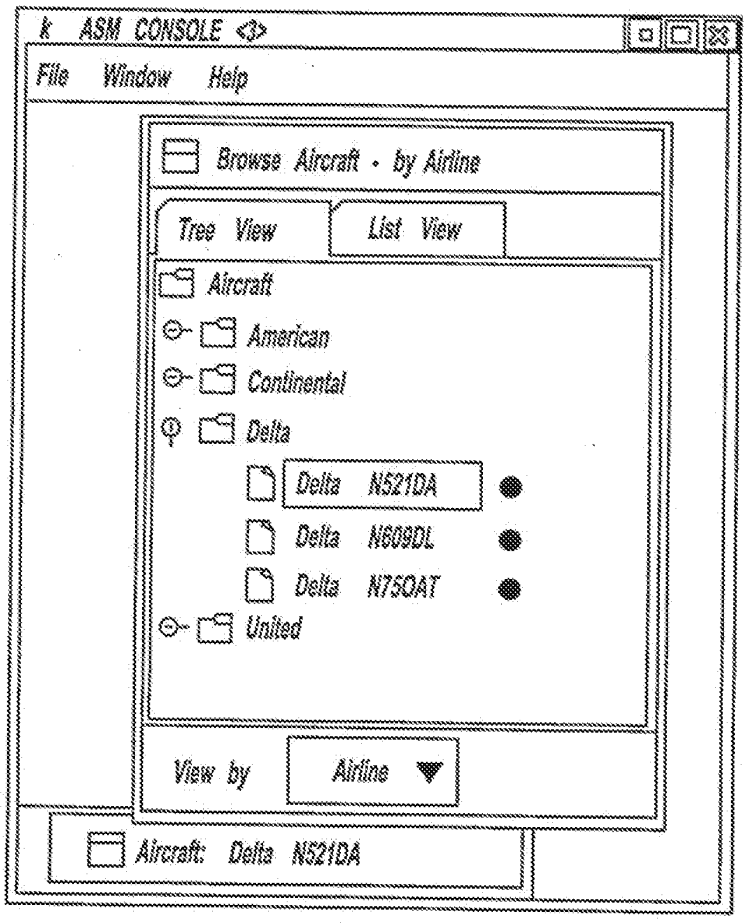

ASM CONSOLE
File
Window
Help
Browse Aircraft - by Airline
Tree View
List View
Aircraft
American
Continental
Delta
Delta N521DA
Delta N609DL
Delta N75OAT
United
View by
Airline
Aircraft: Delta N521DA

Figure - 7

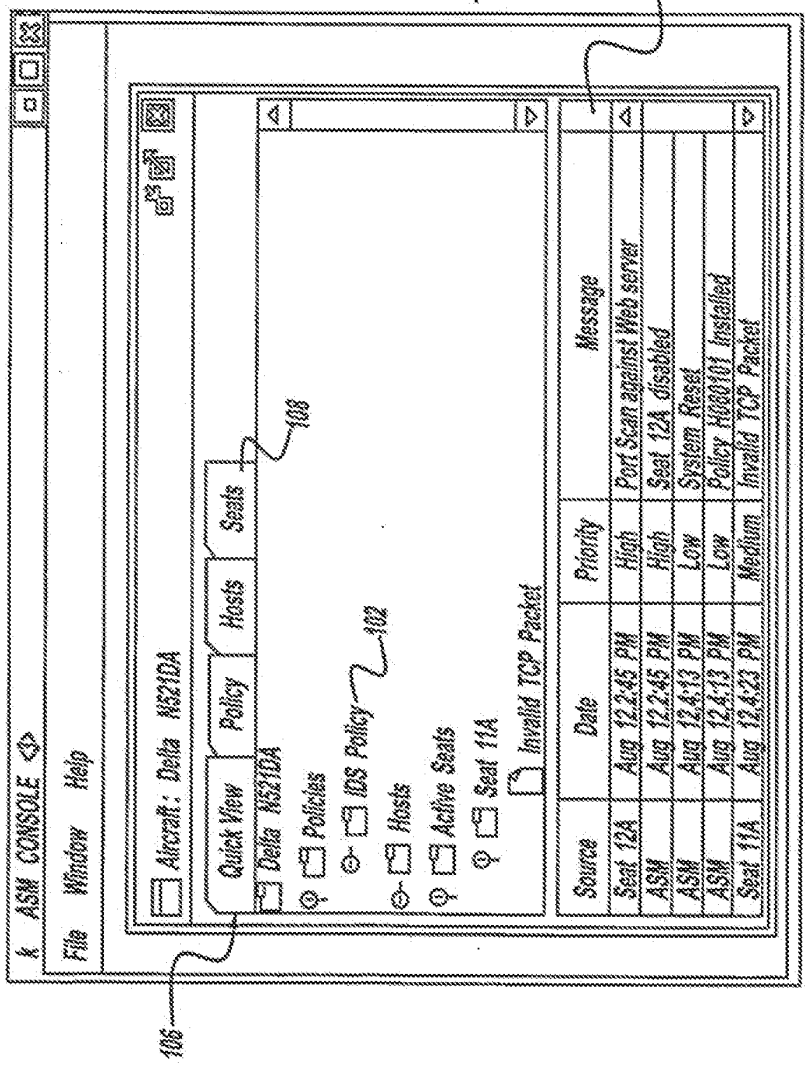

ASM CONSOLE
File Window Help
Aircraft: Delta N521DA
Quick View
Policy
Hosts
Seats
Delta N521DA
Policies
IDS Policy
Hosts
Active Seats
Seat 11A
Invalid TCP Packet
Source
Date
Priority
Message
Seat 12A
Aug 12:2:45 PM
High
Port Scan against Web server
ASM
Aug 12:2:45 PM
High
Seat 12A disabled
ASM
Aug 12:4:13 PM
Low
System Reset
ASM
Aug 12:4:13 PM
Low
Policy H000101 Installed
Seat 11A
Aug 12:4:23 PM
Medium
Invalid TCP Packet
102
104
106
108

Figure - 8

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
27 September 2007 (27.09.2007)

(10) International Publication Number
WO 2007/107701 A2

(51) International Patent Classification:
G06F 21/00 (2006.01)

(21) International Application Number:
PCT/GB2007/000859

(22) International Filing Date: 12 March 2007 (12.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
06251528.3 22 March 2006 (22.03.2006) EP

(71) Applicant *(for all designated States except US)*: **BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY** [GB/GB]; 81 Newgate Street, London Greater London EC1A 7AJ (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **COFTA, Piotr Leon** [PL/GB]; 32 Bugsby Way, Kesgrave, Ipswich Suffolk IP5 2WX (GB).

(74) Agent: **WILLIAMSON, Simeon Paul**; BT GROUP LEGAL INTELLECTUAL PROPERTY DEPARTMENT, PPC5A, BT Centre, 81 Newgate Street, London Greater London EC1A 7AJ (GB).

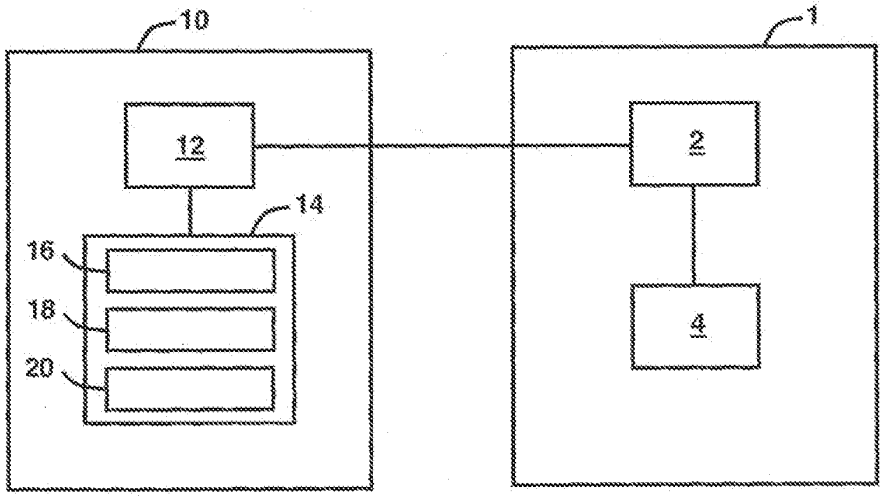
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, ME, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMMUNICATIONS DEVICE MONITORING



(57) Abstract: A method, and devices, for a first communications device (10), e.g. a Subscriber Identity Module (SIM) to monitor a second communications device, e.g. a mobile telephone (1). The method comprises: the SIM (10) sending a program (16) stored at the SIM (10) to the mobile telephone (1); the mobile telephone (1) running the program (16) to provide one or more outcomes; the mobile telephone (1) sending the one or more outcomes to the SIM (10); the SIM (10) comparing the one or more outcomes to one or more correct outcomes (18) stored at the SIM (10); and the SIM (10) evaluating the result of the comparison according to one or more policies (20) stored at the SIM (10).

WO 2007/107701 A2

COMMUNICATIONS DEVICE MONITORING

The present invention relates to monitoring a communications device to verify, determine or otherwise assess the behaviour, condition or integrity of the communications device. The present invention relates in particular, but not exclusively, to monitoring a mobile communications device having a Subscriber Identity Module (SIM) located in the mobile communications device.

In the field of communications devices conventional approaches are provided for a first communications device to monitor a second communications device, in particular so that the first communications device can verify, determine or otherwise assess the behaviour, condition or integrity of the second communications device.

In some conventional approaches, the first communications device observes external signs of activity of the second communications device. In other conventional approaches, the first communications device engages in a pre-defined communications interaction with the second communications device to verify whether the second communications device performs correctly with respect to the interaction. For example, the first communications device may open an Internet connection at set time intervals to determine whether the second communications device is able to accept the Internet connection.

A common type of communication device, the mobile telephone, is becoming increasingly susceptible to outside influences on its integrity, due to increasing use of mobile telephones for data and software download, Internet connection, and so on. For example, mobile telephones are increasingly likely to be subject to malicious influences such as viruses and worms.

US20050033987 describes a mobile terminal that can carry out an integrity check on itself based on a trusted platform concept, for example the Trusted Computing Platform (TCP), using integrity metrics. Data is stored, for example on a SIM in the mobile terminal.

US20040243801 describes a system in which a mobile telephone can perform a self-check, using integrity metrics, in response to a request for verification.

In a first aspect, the present invention provides a method for a first communications device to monitor a second communications device, the method comprising: the first communications device sending a program stored at the first communications device to the second communications device; the second communications device running the program to provide one or more outcomes; the second communications device sending the one or more outcomes to the first communications device; the first communications device comparing the one or more outcomes to one or more correct outcomes stored at the first communications device; and the first communications device evaluating the result of the comparison according to one or more policies stored at the first communications device.

The first communications device may be located in the second communications device.

The first communications device may be a smart card and the second communications device may comprise means for coupling with the smart card.

The first communications device comprises cellular communications system Subscriber Identity Module, SIM, functionality and the second communications device is a mobile communications device that uses the SIM functionality of the first communications device.

The first communications device may not have a capability to run the program.

The first communications device may select the program to send to the second communications device from among a plurality of programs stored at the first communications device.

The first communications device may select randomly or pseudo-randomly the program to send to the second communications device from among a plurality of programs stored at the first communications device.

The method may be such that as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, the first communications device ceases co-operating with the second communications device. Ceasing co-operating may comprise declining to provide cellular communications system authorisation.

The method may be such that as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, the first communications device contacts a network operator of the cellular communications system through the second communications device.

In a further aspect, the present invention provides a first communications device adapted to monitor a second communications device, the first communications device comprising: storage means, the storage means storing a program, correct outcomes of the program, and policies; means adapted to send the program to a second communications device; means adapted to receive one or more outcomes of running the program from the second communications device; means adapted to compare the received one or more outcomes to corresponding correct outcomes of the stored correct outcomes; and means adapted to evaluate the result of the comparison according to one or more of the stored policies.

The first communications device may be adapted to be located in the second communications device.

The first communications device may be a smart card.

The first communications device may comprises cellular communications system Subscriber Identity Module, SIM, functionality.

The first communications device may not have a capability to run the program.

The program may be one of a plurality of programs stored by the storage means; and the first communications device may further comprise means adapted to select the program from among the plurality of programs. The means adapted to select the program from among the plurality of programs may be adapted to perform the selection randomly or pseudo-randomly.

The first communications device may further comprise means for, as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, ceasing co-operating with the second communications device. Ceasing co-operating may comprise declining to provide cellular communications system authorisation.

The first communications device may further comprise means for, as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, contacting a network operator of the cellular communications system through the second communications device.

In a further aspect, the present invention provides a second communications device adapted to be monitored by a first communications device, the second communications device comprising: means adapted to receive a program from a first communications device; means adapted to run the program to provide one or more outcomes; and means adapted to send the one or more outcomes to the first communications device.

The second communications device may further comprise means for locating the first communications device in the second communications device.

The first communications device is a smart card.

The second communications device may be a mobile communications device adapted to use cellular communications system Subscriber Identity Module, SIM, functionality of the first communications device.

In further aspects, the present invention provides a storage medium storing processor-implementable instructions for controlling one or more processors to carry out the method or implement the devices of any of the above aspects.

Thus the present invention provides an active monitoring process in which integrity metrics are not required to be measured.

In certain aspects, the present invention provides an active monitoring process for a mobile device, for example a mobile telephone, by using a SIM, or a card or other module comprising SIM functionality, located in the mobile device, as the monitoring device in the active monitoring process. This advantageously makes use of the SIM's typically greater resistance to outside influences compared to that of the mobile device.

Embodiments of the invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

Figure 1 is a block diagram showing certain functional modules of a mobile device in which a first embodiment of the present invention is implemented;

Figure 2 is a block diagram showing certain functional modules of the mobile device of Figure 1 and certain functional modules of SIM located in the mobile device; and

Figure 3 is a process flowchart showing certain steps carried out by the SIM and the mobile device in an embodiment of a monitoring process.

Figure 1 is a block diagram showing certain functional modules of a mobile device 1 in which a first embodiment of the present invention is implemented. In this embodiment the mobile device 1 is a mobile telephone 1. The mobile telephone 1 comprises a processor 2, hereinafter referred to as the telephone processor 2, and the following modules, each of which are coupled to the processor 2: a radio input/output module 3, a storage medium 4, a user interface 5, a display 6, a loudspeaker 7 and a microphone 8. The storage medium 4 of the mobile telephone 1 is hereinafter referred to

as the telephone storage medium 4. The radio input/output module 3 is further coupled to an antenna 9.

A SIM 10 is shown located in the mobile telephone 1. The SIM 10 is coupled to the telephone processor 2 of the mobile telephone 1. When located in the mobile telephone 1, the SIM 10 is often considered as, or described as, being part of the mobile telephone 1. However, the present embodiment is most readily described by considering the SIM 10 as a separate communications device to the mobile telephone 1. In the present embodiment, the mobile telephone 1 and the SIM 10 operate according to the UMTS specification (Universal Mobile Telecommunication System), as specified in ETSI (European Telecommunications Standards Institute) specifications TS102.221, TS131.101, TS131.102 and other related specifications, except where described otherwise below.

In operation, the mobile telephone 1, under control of the telephone processor 2, and where appropriate using instructions and/or data stored in the telephone storage medium 4, carries out conventional mobile telephone operation according to the UMTS specification, as described in the following overview. Wireless signals are sent to, and received from, base stations of a cellular communications system using the antenna 8, with modulation and demodulation being performed by the radio input/output module 3. The user interface 5, comprising a keyboard and other touch devices enables a user to input instructions and data. The display 6 displays information, such as telephone numbers, network details, images and so on to the user. The loudspeaker 7 and the microphone 8 enable the user to respectively input and hear audio information, e.g. speech during a telephone call. The SIM 10 provides data and cellular communications network connectivity functionality to the mobile telephone 1. This allows the mobile telephone 1 to send and receive short message service (SMS) messages including multi-media messages, partake in telephone calls, access the Internet, download data and software, and so on.

It will be appreciated that the mobile telephone 1 comprises many other conventional components and functions in addition to the functional modules described above. It will also be appreciated that the functional modules described above may exhibit many other conventional functions and details. However, further explanation of any of these additional conventional components, functions and details is not necessary for understanding the present invention, and moreover they are readily implemented by the skilled person.

Figure 2 is a block diagram showing certain of the above described functional modules of the mobile telephone 1 and certain functional modules of the SIM 10. In particular, Figure 2 shows the telephone processor 2 and the telephone storage medium 4 of the mobile telephone 1. As shown in Figure 2, the SIM 10 comprises a processor 12, hereinafter referred to as the SIM processor 12. The SIM processor 12 is coupled to the telephone processor 2. The SIM 12 further comprises a storage medium 14, hereinafter referred to as the SIM storage medium 14. The SIM storage medium 14 is coupled to the SIM processor 12. Operation of the SIM 10 is controlled by the SIM processor 12 in conventional fashion, where appropriate using conventional algorithms, instructions and data (not shown in Figure 2) stored in the SIM storage medium 14.

In this embodiment the SIM storage medium 14 comprises, in addition to the above mentioned conventional algorithms, instructions and data, further algorithms, instructions and data for use in a process of monitoring the mobile telephone 2, which process will be described in more detail below with reference to Figure 3. As shown in Figure 2, the further algorithms, instructions and data for use in the monitoring process comprise test programs 16, correct outcomes 18, and policies 20. In this embodiment the test programs 16 are provided using the SIM Application Toolkit. In particular, proactive commands in the SIM Application Toolkit are used to make the communication between the SIM and the mobile telephone two-way by virtue of the mobile telephone polling the SIM. Further details of the test programs 16, correct outcomes 18, and policies 20 will be described in more detail below.

Figure 3 is a process flowchart showing certain steps carried out by the SIM 10 and the mobile telephone 1 in this embodiment of a monitoring process. At the top of the flowchart, the reference numeral 10 indicates those steps carried out by the SIM 10, and the reference numeral 1indicates those steps carried out by the mobile telephone 1.

At step s2, the SIM 10, more particularly the SIM processor 12, selects one of the test programs 16.

At step s4, the SIM 10 sends the selected test program 16 to the mobile telephone 1.

At step s6, the mobile telephone 1 receives the test program, and under control of the telephone processor 2 stores the test program 16 in the telephone storage medium 4.

At step s8, the SIM 10, more particularly the SIM processor 12, requests the program 16 to be run, by sending a request message to the mobile telephone 1.

A step s10, the mobile telephone 1, more particularly the telephone processor 2, executes the program stored in the telephone storage medium 4.

Meanwhile, at step s12, the SIM 10 awaits the outcomes of the test program 16 being run by the mobile telephone 1.

At step s14, the outcomes of the test program are sent by the mobile telephone 1 to the SIM 10. In this embodiment the outcomes are sent according to instructions contained in the test program 16 itself. This has an advantage that any corrupted program in the mobile telephone need not be involved in the process. Nevertheless, this is not essential, and in other embodiments, the mobile telephone 2 can be pre-programmed with instructions to send any outcomes from test programs to the source of the testing program, i.e. in this case the SIM 10.

At step s16, the SIM 10, more particularly the SIM processor 12, compares the received outcomes with the correct outcomes 18 stored in the SIM storage medium 14.

At step s18, the SIM 10, more particularly the SIM processor 12, determines a decision for the received outcomes, when compared with the correct outcomes 18, according to the policies 20 stored in the SIM storage medium 14. In this embodiment, the policies 20 simply specify that if the comparison step s16 indicates that the received outcomes are the same as the correct outcomes 18 then the SIM 10 continues to co-operate with the mobile telephone 1, whereas if the comparison step s16 indicates that any of the receive outcomes are different to the correct outcomes 18 then the SIM 10 stops co-operating with the mobile telephone 1. Many other policies 20 may be implemented, and other such possibilities will be discussed later below.

At step s20, the SIM 10 implements the decision determined at step s18. Hence, in this embodiment, at step s20, SIM 10 continues to co-operate with the mobile telephone 1 if the received outcomes were the same as the stored correct outcomes 18, whereas SIM 10 stops co-operating with the mobile telephone 1 if the received outcomes were different to the stored correct outcomes 18.

A time out option may be provided such that if the mobile telephone 1 does not return any outcomes within a given time from being instructed to run the test program, then this is in itself considered as a received outcome by the SIM 10, and this outcome is accorded a policy.

It should be noted that certain of the process steps depicted in the flowchart of Figure 3 and described above may be omitted or such process steps may be performed in differing order to that presented above and shown in Figure 3. Furthermore, although all the process steps have, for convenience and ease of understanding, been depicted as

discrete temporally-sequential steps, nevertheless some of the process steps may in fact be performed simultaneously or at least overlapping to some extent temporally. For example, steps s4 and s8 may be performed as one step, i.e. the SIM 10 may send the program to the device and the request to run the program at the same time. Another possibility is omission of the step of the SIM 10 requesting the program to be run, and instead the mobile telephone automatically executing any such received test program.

Thus a process is provided in which a program stored at the SIM 10 is executed by the mobile telephone 1, and the outcome of the program is returned to the SIM to evaluate. Thus a process is provided in which the relatively more secure SIM can effectively probe the status of the relatively less secure mobile telephone. Furthermore, because the test programs are stored at the SIM and not at the mobile telephone (except when due to be executed), any attacker of the mobile telephone cannot ascertain what programs any tampered version of the mobile telephone will need to be able to carry out to mask any changes made to the mobile telephone.

In the above embodiments, the SIM 10 and mobile telephone 1 operate according to the UMTS specification. In other embodiments, the SIM 10 and mobile telephone may instead, or in addition, operate according to other cellular communications specifications, for example Global System for Mobile Communications (GSM).

In the above embodiment the mobile communications device is a mobile telephone. In other embodiments other types of mobile device may be employed, for example a personal digital assistant or other device linked to a mobile radio receiver or module.

Indeed, in other embodiments, the communications device being monitored need not be a mobile device. For example, communications devices which although not portable nevertheless employ a SIM or other smart card, of the type described in the following paragraphs may be the device being monitored.

In the above embodiments, the element performing the monitoring of the mobile telephone is a conventional UMTS SIM adapted by programming using the SIM application toolkit. Also, dual SIM with more than one authentication means can be used. For example, dual SIM with more than one authentication means (for example with two SIM functionalities embedded on the same physical card) can be used. In other embodiments, the SIM functionality used in the above embodiments may be housed and delivered in physical forms other than a conventional SIM card. For example, the functionality may be provided on a Secure Digital (SD) card or a Multi Media Card (MMC). More generally, the device performing the monitoring may be any type of smart card, i.e. a

module comprising processing as well as memory, with the device being monitored being any communications device with means for coupling with the smart card. In each such embodiment, due to the physical location of the SIM or e.g. other smart card in the device being monitored, it can be typically assumed that a required level of trust between the monitoring device and the device being monitored is present. Nevertheless, as a further generalisation, in other embodiments coupling between the SIM or e.g. other smart card with the device being monitored may instead be provided remotely, for example over a Bluetooth wireless link or other local wireless link.

The test programs may be written in any suitable programming language, with the device being monitored being provided with processing means for executing programs in the chosen language. However, preferably the test programs are in a conveniently portable language such as Java. Mobile telephones meeting the UMTS specification typically comprise a Java capability therein. Therefore, such conventional mobile telephones may be simply adapted for use in the above embodiments by providing them with a Java-SIM connection. This may be provided consistent with the JCP (Java Community Process) specification of JSR177 (SATSA): Security and Trust which enables integration of these, and by incorporating a CLDC (Connected Limited Device Configuration) platform. The use of Java is beneficial as it guarantees high portability of the test program between different mobile telephones so that SIM does not have to keep different variants of the same test program for different telephones. Further, the use of SATSA enables the creation of complex test cases while the ability to sign the latest program provides additional security for the mobile device.

It will be appreciated that the test programs may be ones that can be stored in the SIM but are too complex to be have been executed in the SIM. For example, the SIM may have insufficient processing power, whereas the mobile communications device has more processing power and is able to execute the program. This advantageously allows a synergy to be achieved between the security of the SIM and the processing power of the device being monitored, i.e. the relative complexity of the test programs that may be employed.

The test programs may be of any desired complexity. Typically, for higher complexity programs there will be a greater chance of finding problems, whereas for lower complexity programs there will typically be a lower processing and storage burden. One example is, say, to instruct the device being monitored to add three specified memory values and send the sum as an outcome. Another example, say, is to instruct the device being monitored to measure a current and send the measured current value as the

outcome, where the current value is one that indicates that the display, e.g. display 6 in above described mobile telephone 1, is correctly functioning. Another example is the test program can require the device being monitored, e.g. a mobile telephone, to communicate in a specific way with the cellular communications network operator.

The choice of which stored test program to send to the device being monitored can be determined according to any desired criterion according to the requirements of the arrangement under consideration. For example, the test program may be chosen randomly or pseudo-randomly (this option in particular enforces the earlier described advantage of the test programs being stored at the SIM and not at the mobile telephone in relation to any attacker of the mobile telephone therefore not being able to ascertain what programs any tampered version of the mobile telephone will need to be able to carry out to mask any changes made to the mobile telephone). Another possibility is the test program may be selected by sequentially from a sequentially ordered list. For example, yet another possibility is the program may be selected from the stored programs according to some algorithm that takes account of previous outcomes.

In any of the above embodiments, the test programs sent from the SIM or corresponding card or device to the device being monitored may be digitally signed programs, and the device being monitored correspondingly programmed or otherwise set up to grant the necessary rights to the digitally signed programs.

In the above described embodiments, a plurality of test programs are stored at the SIM or corresponding device. However, in other simpler embodiments, only one test program is stored at the SIM and used in the monitoring process.

For the avoidance of doubt, it is noted that the term "program" as used in the above description encompasses similar terms used in the art, such as "application".

The policies described above may be of many different forms. A simple policy is, as described above, for the SIM or corresponding device to cease co-operating with the device being monitored if the outcomes received are not the same as the expected, i.e. correct outcomes. This may include refusing to carry out network authentication, e.g. UMTS or GSM authentication with the network operator. Other policies may be based on more complicated algorithms, for example a weighting or similar operation carried out to determine a significance of one or more differences between received outcomes and correct outcomes. The policies may be such that resulting actions may further include contacting the network operator to report a discrepancy. This may be over a secure communication link provided via the device being monitored, for example when the device being monitored is a mobile telephone, the secure communication link may comprise

sending an SMS message or call to the network operator. The policies may be hard-coded, or may updatable in some manner. For example, they may be altered by user input. Another possibility is that the network operator may use a secure communication through the device being monitored to communicate with the SIM. For example, the network operator may change settings in the SIM memory, thereby changing the selection of test programs, policies, correct outcomes and so on.

In the above embodiments, a plurality of policies are stored at the SIM. However, in simpler embodiments, only one policy is stored and used, e.g. a policy of ceasing co-operation with the device being monitored in the event of any received outcome not matching the corresponding stored correct outcome.

In the above embodiments, there a plurality of outcomes are provided when the test program is run. However, in simpler embodiments, there may be just one outcome when the test program is run.

In the above embodiments, the test programs are provided using the SIM Application Toolkit. In particular, proactive commands in the SIM Application Toolkit are used to make the communication between the SIM and the mobile telephone two-way by virtue of the mobile telephone polling the SIM. In other embodiments, other functionality may be added to a conventional SIM (or corresponding device or card) and/or the mobile telephone or other device being monitored such as to provide that the SIM may push the test programs to the mobile telephone and/or the mobile telephone may pull the test programs from the SIM. One such possibility is that a conventional smartphone (e.g. in a form of PDA or a handheld device) equipped with an open operating system such as Symbian or Windows Mobile as typically used in mobile telephones under UMTS or GSM, may be adapted as required.

More generally, the above embodiments may be implemented by providing new designs of communications device, e.g. mobile telephone and/or SIM card, or by configuring or adapting previous designs of device. The described functionality may be provided by hardware, firmware, software, or any combination of these For example existing processors/storage medium may be programmed (for example with data stored at the described storage medium) to provide the above described processes. As such the described processes may be implemented by a processor implementing processor-implementable instructions stored at the storage medium. The storage medium may be any suitable storage medium, such as computer memory, ROM, PROM etc. Any described processor may comprise plural processors.

CLAIMS

1. A method for a first communications device to monitor a second communications device, the method comprising:

the first communications device sending a program stored at the first communications device to the second communications device;

the second communications device running the program to provide one or more outcomes;

the second communications device sending the one or more outcomes to the first communications device;

the first communications device comparing the one or more outcomes to one or more correct outcomes stored at the first communications device; and

the first communications device evaluating the result of the comparison according to one or more policies stored at the first communications device.

2. A method according to claim 1 wherein the first communications device is located in the second communications device.

3. A method according to claim 1 or 2 wherein the first communications device is a smart card and the second communications device comprises means for coupling with the smart card.

4. A method according to any of claims 1, 2 or 3, wherein the first communications device comprises cellular communications system Subscriber Identity Module, SIM, functionality and the second communications device is a mobile communications device that uses the SIM functionality of the first communications device.

5. A method according to any of claims 1 to 4, wherein the first communications device does not have a capability to run the program.

6. A method according to any of claims 1 to 5, wherein the first communications device selects the program to send to the second communications device from among a plurality of programs stored at the first communications device.

7. A method according to claim 6, wherein the first communications device selects randomly or pseudo-randomly the program to send to the second communications device from among the plurality of programs stored at the first communications device.

8. A method according to any of claims 1 to 7, wherein as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, the first communications device ceases co-operating with the second communications device.

9. A method according to claim 8 when dependent on claim 4, wherein ceasing co-operating comprises declining to provide cellular communications system authorisation.

10. A method according to claim 4, 8 or 9, wherein as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, the first communications device contacts a network operator of the cellular communications system through the second communications device.

11. A first communications device adapted to monitor a second communications device, the first communications device comprising:

storage means, the storage means storing a program, correct outcomes of the program, and policies;

means adapted to send the program to a second communications device;

means adapted to receive one or more outcomes of running the program from the second communications device;

means adapted to compare the received one or more outcomes to corresponding correct outcomes of the stored correct outcomes; and

means adapted to evaluate the result of the comparison according to one or more of the stored policies.

12. A first communications device according to claim 11, wherein the first communications device is adapted to be located in the second communications device.

13. A first communications device according to claim 11 or 12, wherein the first communications device is a smart card.

14. A first communications device according to any of claims 11 to 13, wherein the first communications device further comprises cellular communications system Subscriber Identity Module, SIM, functionality.

15. A first communications device according to any of claims 11 to 14, wherein the first communications device does not have a capability to run the program.

16. A first communications device according to any of claims 11 to 15, wherein the program is one of a plurality of programs stored by the storage means; and the first communications device further comprises means adapted to select the program from among the plurality of programs.

17. A first communications device according to claim 16, wherein the means adapted to select the program from among the plurality of programs are adapted to perform the selection randomly or pseudo-randomly.

18. A first communications device according to any of claims 11 to 17, further comprising means for, as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, ceasing co-operating with the second communications device.

19. A first communications device according to claim 18 when dependent on claim 14, wherein ceasing co-operating comprises declining to provide cellular communications system authorisation.

20. A first communications device according to claim 14, 18 or 19, further comprising means for, as a result of the evaluating step, when one or more received outcomes is not the same as the correct outcomes, contacting a network operator of the cellular communications system through the second communications device.

21. A second communications device adapted to be monitored by a first communications device, the second communications device comprising:

means adapted to receive a program from a first communications device;

means adapted to run the program to provide one or more outcomes; and

means adapted to send the one or more outcomes to the first communications device.

22. A second communications device according to claim 21, further comprising means for locating the first communications device in the second communications device.

23. A second communications device according to claim 21 or 22, wherein the first communications device is a smart card.

24. A second communications device according to any of claims 21 to 23, wherein the second communications device is a mobile communications device adapted to use cellular communications system Subscriber Identity Module, SIM, functionality of the first communications device.

25. A storage medium storing processor-implementable instructions for controlling one or more processors to carry out the method of any of claims 1 to 10.

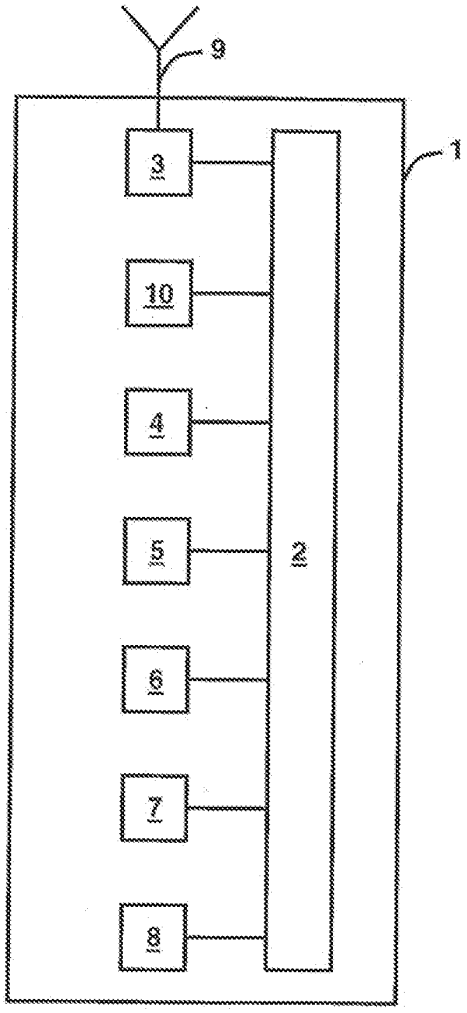


FIG. 1

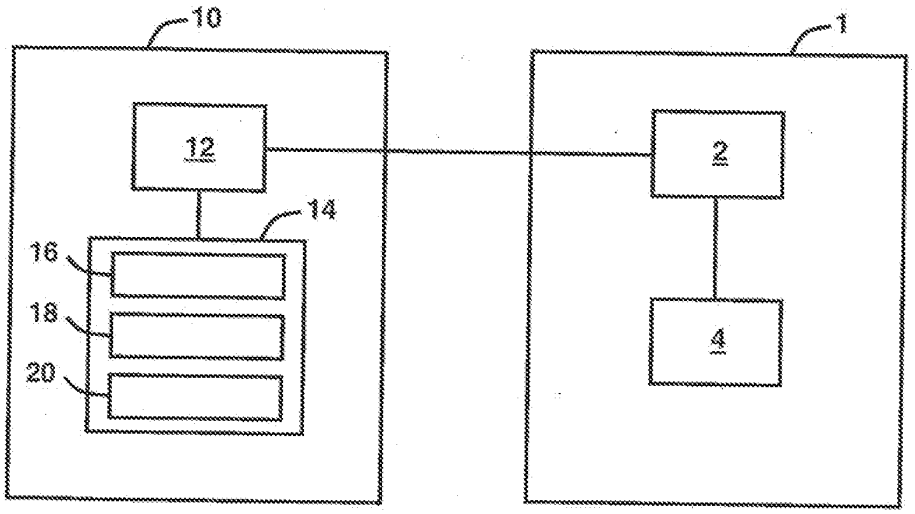


FIG. 2

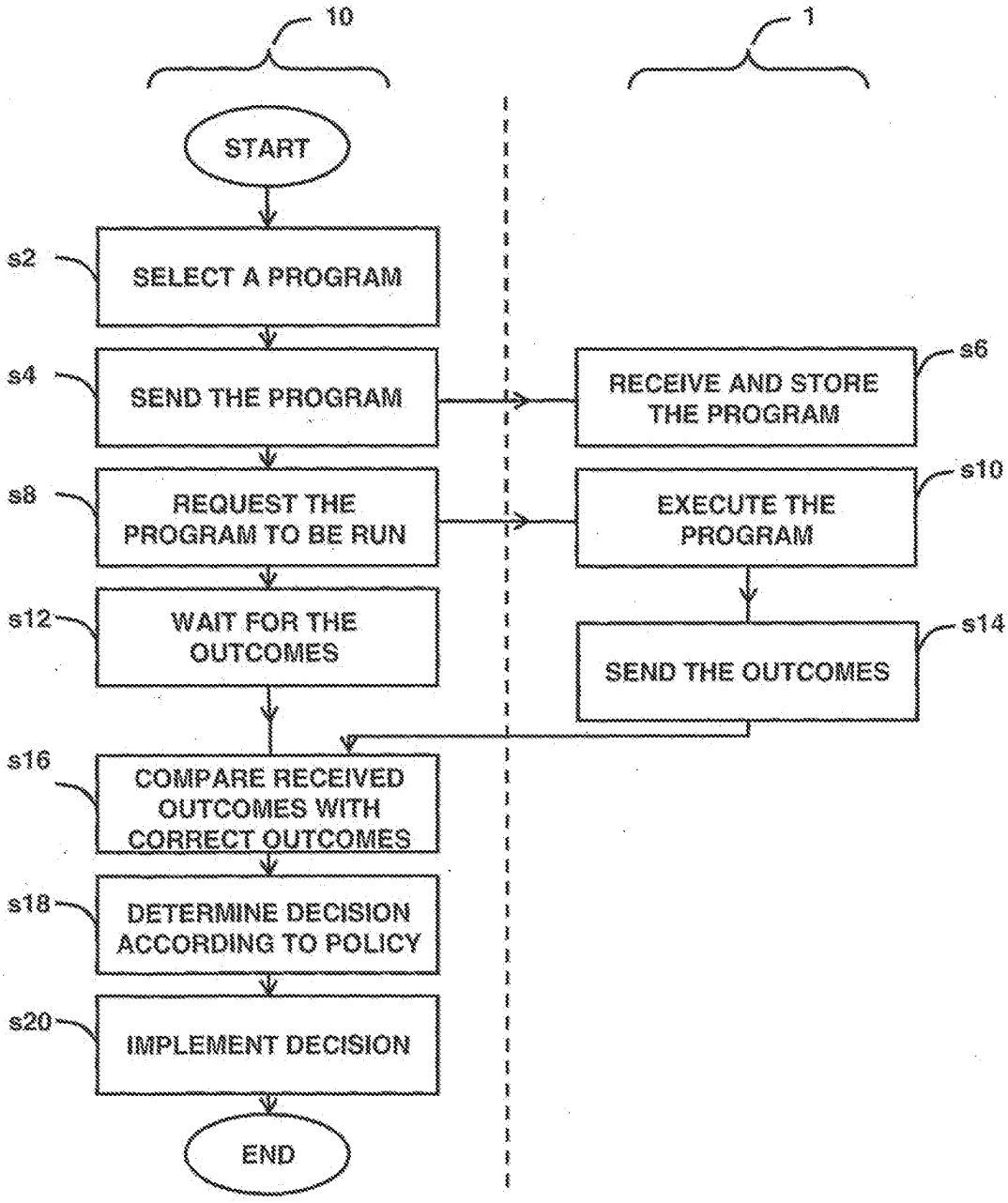

10
1
START
s2
SELECT A PROGRAM
s4
SEND THE PROGRAM
s6
RECEIVE AND STORE THE PROGRAM
s8
REQUEST THE PROGRAM TO BE RUN
s10
EXECUTE THE PROGRAM
s12
WAIT FOR THE OUTCOMES
s14
SEND THE OUTCOMES
s16
COMPARE RECEIVED OUTCOMES WITH CORRECT OUTCOMES
s18
DETERMINE DECISION ACCORDING TO POLICY
s20
IMPLEMENT DECISION
END

FIG. 3

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 463 238 A1

(12) DEMANDE DE BREVET EUROPEEN

(43) Date de publication:
29.09.2004 Bulletin 2004/40

(51) Int Cl.?: **H04L 12/24**, H04L 12/26

(21) Numéro de dépôt: **04290706.3**

(22) Date de dépôt: **15.03.2004**

(84) Etats contractants désignés:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PL PT RO SE SI SK TR
Etats d'extension désignés:
AL HR LT LV MK

(30) Priorité: **28.03.2003 FR 0303835**

(71) Demandeur: **ALCATEL**
75008 Paris (FR)

(72) Inventeurs:
- **Marilly, Emmanuel**
  92160 Antony (FR)
- **Betge-Brezetz, Stéphane**
  75015 Paris (FR)
- **Martinot, Olivier**
  91210 Draveil (FR)
- **Chevanne, Michel**
  92140 Clamart (FR)
- **Delegue, Gérard**
  94230 Cachan (FR)

(74) Mandataire: **Chaffraix, Sylvain et al**
Compagnie Financière Alcatel
Département de Propriété Industrielle,
5, rue Noel Pons
92734 Nanterre Cedex (FR)

(54) Dispositif de gestion locale de procédés d'assurance pour un équipement de réseau de communications

(57) Un dispositif embarqué (D) est dédié à la gestion locale d'assurance pour un équipement de réseau (NE) d'un réseau de communications (N) équipé d'un système de gestion de réseau (NMS), ledit équipement (NE) présentant une configuration choisie et comportant des moyens (MM) de mesure de valeurs de paramètre (s) du réseau et une base d'informations de gestion (MIB) stockant des données de gestion représentatives des valeurs mesurées. Le dispositif (D) comprend des moyens de gestion (MAE) chargés d'adapter la configuration de l'équipement (NE) en fonction d'au moins les données de gestion stockées dans la base d'informations de gestion (MIB) et de règles, dites d'assurance, choisies.

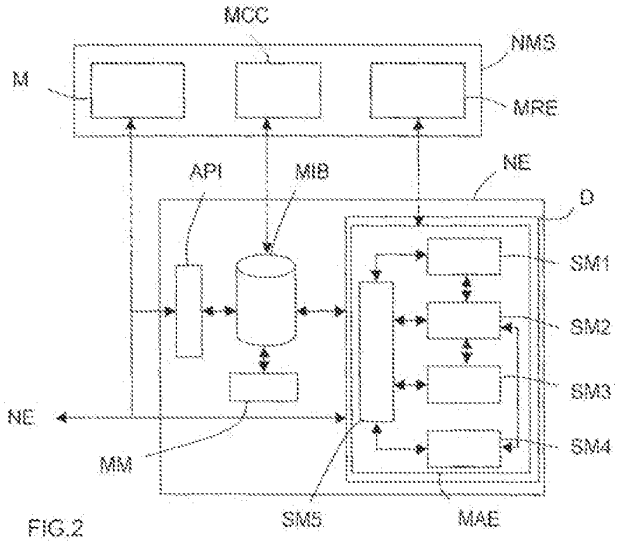


FIG.2

EP 1 463 238 A1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] L'invention concerne les réseaux de communications, et plus particulièrement ceux offrant une assurance de continuité de service(s).

[0002] De nombreux réseaux de communications mettent en oeuvre des procédés d'assurance qui contrôlent l'exécution d'activités de maintenance préventives et réactives destinées à assurer à leurs clients la continuité de la disponibilité de services, définis en termes de qualité de service (QoS). A cet effet, les réseaux disposent d'un système de gestion de réseau (ou NMS pour « Network Management System ») qui analyse continuellement l'état des ressources et les performances de manière à anticiper ou détecter des problèmes ou pannes et à déterminer les actions à entreprendre de sorte que les clients ne soient pas pénalisés.

[0003] Une telle analyse centralisée repose sur la collection d'un très grand nombre de données d'information et d'alarmes auprès des nombreux équipements de réseau, chargés d'effectuer des mesures de valeurs de paramètre(s) du réseau, et plus précisément auprès de leurs bases d'informations de gestion (ou MIB pour « Management Information Base ») chargées de stocker les données de gestion représentatives des valeurs mesurées. Une fois collectées, par exemple via un collecteur SNMP, les données sont stockées dans une base de données d'archivage, par exemple de type Oracle, et alimentent des modules de calcul qui mettent en oeuvre des formules ou équations prédéfinies.

[0004] Une telle solution est par exemple décrite dans l'article « *Real-Time Performance Monitoring and Anomaly Detection in the Internet: and Adaptative, Objective-Driven, Mix-and-Match Approach* » de L. Lawrence Ho, Christopher J. Macey and Ronald G. Hiller.

[0005] Un tel procédé d'assurance centralisé requiert l'emploi de bases de données de très grandes capacités, nécessite des temps de calcul importants et consomme de la bande passante (lors de la collecte), ce qui nuit aux performances du réseau.

[0006] Certaines tentatives ont été faites vers l'utilisation d'agents embarqués dans des éléments réseaux, mais les aspects décisionnels demeurent dans un organe centralisé et présentent les mêmes défauts que la solution précédente. De plus, ces tentatives ne concernent pas l'assurance.

[0007] Des exemples de telles tentatives peuvent être trouvées dans les brevets américains US 6 272 537 et US 6 085.

[0008] L'invention a donc pour but d'améliorer la situation.

[0009] Elle propose à cet effet un dispositif de gestion locale d'assurance pour un équipement de réseau de communications présentant une configuration choisie et comportant des moyens de mesure de valeurs de paramètre(s) du réseau et une base d'informations de gestion (MIB) stockant des données de gestion représentatives des valeurs mesurées.

[0010] Ce dispositif se caractérise par le fait qu'il comporte des moyens de gestion chargés d'adapter la configuration de l'équipement dans lequel il est implanté (ou embarqué) en fonction, notamment, des données de gestion stockées dans sa base d'informations de gestion et de règles, dites d'assurance, choisies, définissant une politique locale d'assurance.

[0011] En d'autres termes, l'invention consiste à intégrer dans des équipements (ou noeuds) du réseau un dispositif d'assurance embarqué de sorte qu'ils puissent assurer localement au moins une partie du procédé d'assurance géré par le NMS.

[0012] Le dispositif d'assurance embarqué, selon l'invention, pourra comporter des caractéristiques complémentaires qui pourront être prises séparément et/ou en combinaison, et en particulier :

- des moyens de gestion capables d'adapter la configuration de leur équipement en fonction de données d'informations provenant d'au moins un autre équipement de réseau,
- une adaptation consistant en une modification de la politique de mesure de paramètre(s) et/ou une modification de la politique d'envoi de rapport (ou « reporting ») au NMS,
- une adaptation consistant en une modification du mode de fonctionnement de l'équipement,
- des moyens de gestion comprenant des moyens d'analyse de paramètres du réseau capables de délivrer des données d'information représentatives de l'évolution temporelle, sur un intervalle choisi, de certaines valeurs de paramètre(s) du réseau (comme par exemple le trafic ou l'état d'un routeur) qui sont stockées dans la MIB. Dans ce cas, les moyens d'analyse peuvent être agencés de manière à délivrer des données d'information représentatives d'une analyse de tendance et/ou d'une analyse de profils ou de signatures (ou « baselines ») et/ou d'une analyse de discontinuité et/ou d'une agrégation de valeurs de paramètres du réseau. Par ailleurs, ces moyens d'analyse peuvent être configurables, notamment dans le but de mettre en oeuvre de nouveaux calculs, relatifs aux paramètres du réseau, transmis par le NMS,
- des moyens de gestion comprenant des moyens d'alarme chargés de déclencher l'envoi d'une alarme et/ou de données d'information au NMS et/ou à au moins un autre équipement de réseau, en fonction de règles d'assurance choisies. De telles données d'information et alarmes sont par exemple représentatives de résultat(s) d'analyse(s) ou d'agrégation(s) de données effectuée(s) par les moyens d'analyse et/ou de valeur(s) de paramètre(s) du réseau stockées dans la MIB. Par ailleurs, les moyens d'alarme peuvent être configurables,
- des moyens de gestion comprenant des moyens d'observation du réseau définissant un agent de mesure de flux de type « bout-en-bout » (ou « end-

to-end »), et délivrant des données d'informations. Dans ce cas, les moyens d'observation du réseau peuvent être configurables,

- des moyens de gestion comprenant des moyens de gestion d'accords de niveau de service (ou SLA pour « Service Level Agreement ») délivrant des données d'informations. Dans ce cas, les moyens de gestion d'accords de niveau de service peuvent être configurables,
- des moyens de gestion capables d'effectuer des tests, par exemple de mesure(s) active(s) ou d'aide au diagnostic. Ces tests peuvent notamment permettre de faire remonter une information agrégée ou le résultat du test (par exemple le temps pris pour une opération de type TCP) au niveau de la couche NMS,
- des moyens de gestion capables d'assurer leur gestion en fonction d'un calendrier, de sorte que certaines (ou toutes les) fonctions d'assurance ou tests ne soient effectuées que certains jours de la semaine ou du mois ou à certaines heures, par exemple,
- des moyens de gestion comprenant des moyens de contrôle chargés de gérer le fonctionnement (et notamment la configuration) des moyens d'analyse, des moyens d'alarme, des moyens d'observation du réseau et des moyens de gestion d'accords de niveau de service, en fonction des règles d'assurance choisies et des données d'informations qu'ils délivrent. De tels moyens de contrôle peuvent se présenter sous la forme d'un moteur de règles (ou « rule engine ») dans lequel sont stockées les règles d'assurance choisies. Par ailleurs, ces moyens de contrôle peuvent être configurables,
- des moyens de gestion configurables,
- des moyens de gestion et/ou des moyens d'analyse et/ou des moyens et/ou des moyens de contrôle et/ou des moyens d'observation du réseau et/ou des moyens de gestion d'accords de niveau de service agencés de manière à pouvoir être configurés par le NMS via une interface de langage de programmation (ou API pour « Application Programming Interface ») et éventuellement via la MIB. Dans ce cas, la MIB peut être configurée ou programmée par le NMS via l'API,
- des moyens de gestion et/ou des moyens d'analyse et/ou des moyens d'alarme et/ou des moyens d'observation du réseau et/ou des moyens de gestion d'accords de niveau de service et/ou des moyens de contrôle agencés de manière à pouvoir être configurés par le NMS via des commandes dédiées, telles que des « Command Line Interface » (ou CLI).

[0013] L'invention concerne également un équipement de réseau muni d'un dispositif du type de celui présenté ci-avant. L'invention convient plus particulièrement, bien que de façon non restrictive, aux équipements de réseau tels que les routeurs, les commutateurs (ou « switchs ») et les pare-feux (ou « firewalls »).

[0014] L'invention propose également un réseau de communications équipé d'un système de gestion de réseau (NMS) et d'une multiplicité d'équipements de réseau du type de celui présenté ci-avant. Dans un mode de réalisation de type distribué, chaque équipement de réseau peut être agencé de manière à délivrer au système de gestion de réseau des alarmes et/ou des données d'information de types différents.

[0015] L'invention peut notamment être mise en œuvre dans toutes les technologies réseaux qui doivent être gérées, et notamment dans les réseaux de transmission (par exemple de type WDM, SONET, SDH), de données (par exemple de type Internet-IP ou ATM) ou de voix (par exemple de type classique, mobile ou NGN).

[0016] D'autres caractéristiques et avantages de l'invention apparaîtront à l'examen de la description détaillée ci-après, et des dessins annexés, sur lesquels :

- la figure 1 illustre de façon schématique un réseau de communications comprenant un système de gestion de réseau (NMS) et des équipements de réseau selon l'invention, et
- la figure 2 illustre de façon schématique un exemple de réalisation d'un équipement de réseau selon l'invention.

[0017] Les dessins annexés pourront non seulement servir à compléter l'invention, mais aussi contribuer à sa définition, le cas échéant.

[0018] L'invention a pour objet de permettre la gestion locale, par des équipements de réseau, d'une partie du procédé d'assurance géré par le système de gestion de réseau (ou NMS pour « Network Management System ») d'un réseau de communications.

[0019] Comme cela est illustré sur la figure 1, un réseau de communications N comporte schématiquement une multiplicité d'équipements de réseau NEi (ici i = 1, 2, mais il peut prendre n'importe quelle valeur), comme par exemple des serveurs équipés d'un pare-feu ou (« firewall »), des commutateurs (ou « switchs »), des routeurs périphériques (ou « edge routers ») ou des routeurs de coeur (ou « core routers »), pouvant échanger des données, selon un protocole de gestion de réseau, avec un système de gestion de réseau, et notamment avec son serveur de gestion NMS.

[0020] Chaque équipement NEi comporte classiquement une base d'informations de gestion MIB (« Management Information Base »), également appelée base d'instances d'objets. Chaque MIB stocke des données de gestion représentatives de valeurs de champs d'information qui caractérisent l'équipement NEi associé. Certains de ces champs d'information désignent des paramètres du réseau dont les valeurs instantanées sont mesurées par des sondes MM implantées dans, ou gérées par, l'équipement NEi. Par ailleurs,

chaque MIB est associée à une définition de base d'informations de gestion (non représentée), également appelée définition de MIB, stockée dans le système de gestion de réseau et accessible au serveur de gestion NMS.

[0021] Dans ce qui suit, on considère à titre d'exemple non limitatif que le réseau de communications N est de type Internet (IP) et que le protocole de gestion du réseau est le protocole SNMP (pour « Simple Network Management Protocol » RFC 2571-2580). Bien entendu, l'invention s'applique à d'autres types de réseau, comme par exemple aux réseaux de transmission de type WDM, SONET ou SDH, de données de type ATM, ou de voix de type classique, mobile ou NGN, et à d'autres protocoles de gestion de réseau, comme par exemple TL1, CORBA ou CMISE/CMIP.

[0022] Par ailleurs, on considère dans ce qui suit que les équipements de réseau NEi sont des routeurs périphériques (ou de bordure) agissant en tant qu'agents d'observation de flux de type « bout-en-bout » (ou « end-to-end »). Ces routeurs sont généralement placés aux frontières du réseau et permettent l'établissement de liens avec d'autres réseaux. Ces agents sont généralement destinés à effectuer des mesures de bout en bout pour des problèmes d'échelle (ou « scalability »). Mais, bien entendu, l'invention n'est pas limitée à ce seul type d'équipement de réseau. Elle concerne tous les équipements de réseau capables d'échanger des données de gestion avec le NMS, et notamment les routeurs périphériques qui n'agissent pas en tant qu'observateurs, les routeurs de coeur, les commutateurs et les pare-feux.

[0023] Chaque équipement de réseau NEi est généralement configuré spécifiquement de manière à transmettre au NMS (ou couche NMS) des données d'information représentatives de mesures qu'il a effectuées avec sa ou ses sondes MM et qu'il a stockées dans sa MIB, ainsi que des alarmes signalant des problèmes ou des pannes. La configuration est généralement définie par une ou plusieurs politiques. Par exemple, une première politique est relative aux conditions d'envoi (ou « reporting ») de valeurs mesurées par l'équipement NEi à l'aide de ses sondes MM. Une deuxième politique peut concerner la ou les mesures de paramètre(s) que l'équipement NEi doit effectuer avec sa ou ses sondes MM. Une troisième politique peut concerner le mode de fonctionnement de l'équipement NEi. Ces politiques sont habituellement définies pour chaque équipement NEi par un module de configuration MC du NMS. Elles sont transmises aux équipements NEi, via le réseau N, à l'aide du protocole de gestion (ici SNMP).

[0024] Mais, le NMS, et notamment son module de collection MCC (ou « collector »), peut également, lorsqu'il l'estime nécessaire, adresser à un équipement de réseau NEi des requêtes pour qu'il lui transmette les valeurs de certains des champs d'information qui sont stockés dans sa MIB. C'est ce que l'homme de l'art appelle, en anglais, le « polling ».

[0025] Comme cela est mieux illustré sur la figure 2, l'invention propose un dispositif de gestion locale d'assurance D destiné à être soit implanté (ou embarqué) dans un équipement NEi, soit intégré dans un boîtier externe raccordé directement à un équipement NEi. Dans ce qui suit, on considère que le dispositif D est embarqué dans un équipement NEi.

[0026] Ce dispositif D selon l'invention comporte un module d'assurance MAE (ou module de gestion locale d'assurance) chargé d'adapter la configuration de l'équipement NEi, dans lequel il est embarqué, en fonction de règles, dites d'assurance, choisies, des données de gestion stockées dans sa MIB, et d'éventuelles données d'information qu'il reçoit des autres équipements de réseau.

[0027] On entend ici par « adapter », le fait de modifier une politique de mesure de paramètre(s) (voire mettre en place une nouvelle politique de mesure(s) ou reconfigurer la politique de mesures(s)) et/ou une politique d'envoi de rapport (ou « reporting ») au NMS et/ou le mode de fonctionnement de l'équipement NEi, initialement défini(s) par le module de configuration MC du NMS.

[0028] Par ailleurs, on entend par « règles d'assurance », des règles définissant la politique d'assurance locale de l'équipement compte tenu des valeurs prises par certains paramètres du réseau, stockées dans sa MIB, ou d'événements extérieurs, tels qu'une action de la couche NMS ou d'un autre routeur.

[0029] Le module d'assurance embarqué MAE permet ainsi d'embarquer dans un équipement NEi certaines fonctions d'assurance de manière à le rendre « plus intelligent », à limiter le nombre d'informations qui transitent dans le réseau, et éventuellement à permettre sa reconfiguration partielle, en fonction de règles d'assurance.

[0030] Afin de gérer localement la politique d'assurance, le module d'assurance MAE comporte préférentiellement un premier sous-module SM1 chargé d'appliquer un ou plusieurs traitements aux valeurs de certains champs d'information stockés dans la MIB. Plus précisément, ces traitements consistent à appliquer aux valeurs des champs d'information choisis, en fonction de règles d'assurance dédiées, des formules ou des équations, définies par le module de configuration MC du NMS, afin de délivrer des données d'information.

[0031] Par exemple, le premier sous-module SM1 est chargé d'effectuer des analyses de tendance sur un intervalle choisi (ou suivant un calendrier prédéfini), utiles aux anticipations de problèmes ou de pannes. Il peut notamment signaler que le pourcentage d'utilisation d'un CPU d'un routeur va atteindre un seuil critique dans une heure, ou que la bande passante d'un LSP va dépasser le seuil choisi dans deux heures.

[0032] Il peut être également chargé d'effectuer des analyses de profils ou de signatures (ou « baselines »). Cela est notamment utile lorsque l'on cherche, par exemple, à reconnaître ou identifier un profil caractéris-

tique d'utilisation, tel qu'une attaque spécifique d'un routeur, ou un mode de facturation (ou « billing »).

[0033] Il peut être également chargé d'effectuer des analyses de discontinuité de flux ou de trafic. Par exemple, en cas de surcharge d'un réseau, des paquets sont supprimés (ou « dropped »), si bien que la détection d'une discontinuité sur le paramètre de suppression (de 0 à x) permet d'informer que l'on a effectivement des paquets supprimés.

[0034] Il peut encore être chargé d'effectuer des agrégations de valeurs de paramètres du réseau, par exemple pour appliquer des formules du module de collection de données MCC dans le routeur ou pour calculer une moyenne sur une période donnée (par exemple la bande passante moyenne utilisée au cours de la dernière semaine). Dans ce cas, les valeurs agrégées peuvent être stockées dans la MIB et en être extraites sur requête de polling issue du module de collection MCC du NMS.

[0035] A titre d'exemple illustratif, le premier sous-module SM1 peut être chargé d'estimer l'évolution prévisionnelle (ou tendance) du nombre de paquets abandonnés (« dropped »), ainsi que l'évolution prévisionnelle de la bande passante utilisée et/ou de la charge de calcul (CPU) de l'équipement NEi. Il peut être également agencé pour effectuer des tests, par exemple de mesure(s) active(s) ou d'aide au diagnostic. Ces tests peuvent notamment permettre de faire remonter une information agrégée ou le résultat du test (par exemple le temps pris pour une opération de type TCP) au niveau de la couche NMS.

[0036] Préférentiellement, le premier sous-module SM1 délivre des données d'informations représentatives du résultat de ses analyses (ou estimations) à un (cinquième) sous-module de gestion (SM5), qui décide des actions à entreprendre en fonction de la politique d'assurance locale de son équipement NEi.

[0037] Le module d'assurance MAE comporte également de préférence un deuxième sous-module SM2 chargé de déclencher l'envoi d'alarme(s) et/ou de données d'information (ou rapports) sur ordre du cinquième sous-module SM5, en fonction de règles d'assurance qui lui sont dédiées.

[0038] Mais, on peut également envisager que le premier sous-module SM1 alimente le deuxième sous-module SM2 en données d'information, de sorte qu'il leur applique certaines des règles d'assurance et décide soit de leur envoi, soit de l'émission d'une alarme, à destination du NMS, et plus précisément d'un module de réception d'évènements MRE, et/ou d'au moins un autre équipement de réseau.

[0039] Dans l'exemple illustré d'un routeur périphérique de type « bout-en-bout », le module d'assurance MAE comporte également de préférence un troisième sous-module SM3 chargé d'observer le réseau, et notamment le trafic de type « bout-en-bout », en fonction de règles d'assurance qui lui sont dédiées. Par exemple, on dispose d'une sonde embarquée, pour effectuer les mesures de bout-en-bout. Cette sonde embarquée est alors configurée de manière à déterminer le flux à étudier et les résultats des mesures alimentent des règles d'assurance dédiées. Par exemple, l'une de ces règles peut être : « si entre deux routeurs périphériques on mesure deux flux différents, alors on agrége la mesure des deux flux pour limiter le calcul ».

[0040] Dans ce cas, le troisième sous-module SM3 définit ce que l'homme de l'art appelle un agent de mesure de flux « bout-en-bout » (ou « end-to-end »).

[0041] Préférentiellement, le troisième sous-module SM3 délivre des données d'informations représentatives du résultat de ses observations (ou mesures) au (cinquième) sous-module de gestion (SM5), qui décide des actions à entreprendre en fonction de la politique d'assurance locale de son équipement NEi. Mais, on pourrait également envisager que le troisième sous-module SM3 alimente le deuxième sous-module SM2 en données d'informations, de sorte qu'il leur applique certaines des règles d'assurance et décide soit de leur envoi, soit de l'émission d'une alarme, à destination du NMS, et plus précisément d'un module de réception d'évènements MRE, et/ou d'au moins un autre équipement de réseau.

[0042] Par ailleurs, et comme illustré sur la figure 2, le module d'assurance MAE peut comporter un quatrième sous-module SM4 chargé de gérer localement certains accords de niveau de service (ou SLA pour « Service Level Agreement »), en fonction de règles d'assurance qui lui sont dédiées. Par exemple, le sous-module SM4 connait l'accord de niveau de service (SLA) d'un client, de sorte qu'il peut être chargé de vérifier si ce SLA est bien respecté, et générer une alarme (ou données d'informations) si ce n'est pas le cas (ou si la situation se dégrade).

[0043] Préférentiellement, le quatrième sous-module SM4 délivre des données d'informations représentatives de comptes-rendus de gestion au cinquième sous-module SM5, qui décide des actions à entreprendre en fonction de la politique d'assurance locale de son équipement NEi. Mais, on pourrait également envisager que le quatrième sous-module SM4 alimente le deuxième sous-module SM2 en données d'information, de sorte qu'il leur applique certaines des règles d'assurance et décide soit de leur envoi, soit de l'émission d'une alarme, à destination du NMS, et plus précisément d'un module de réception d'évènements MRE, et/ou d'au moins un autre équipement de réseau.

[0044] Comme indiqué plusieurs fois ci-avant, le module d'assurance MAE comporte un cinquième sous-module SM5 (ou module de contrôle) chargé de contrôler le fonctionnement des autres sous-modules (ici SM1 à SM4). Plus préférentiellement, ce cinquième sous-module SM5 se présente sous la forme d'un moteur de règles (ou « rule engine ») dans lequel sont stockées toutes les règles d'assurance dédiées aux autres sous-modules. Il s'agit par exemple d'un « Java expert System Shell ».

[0045] Alors que chaque sous-module SM1 à SM4 effectue des calculs et/ou des analyses et/ou des rapports définis à partir de règles d'assurance dédiées, le cinquième sous-module SM5 gère préférentiellement l'ensemble desdites règles d'assurance dédiées. Par « gérer l'ensemble des règles », on entend ici gérer les conditions et les actions définies par les différentes règles.

[0046] Le cinquième sous-module SM5 a plus précisément deux fonctions principales

[0047] Une première fonction consiste à ordonner au deuxième sous-module SM2 de générer une alarme, ou d'adresser un rapport, à destination de la couche NMS (ou d'un autre équipement du réseau), parce qu'en appliquant une règle d'assurance à des données délivrées par l'un des autres sous-modules SM1 à SM4, il s'est aperçu que la condition correspondante n'était pas satisfaite.

[0048] Par exemple, il ordonne la génération d'une alarme parce que les données d'analyse délivrées par le premier sous-module SM1 ne satisfont pas la condition qui leur est associée par la politique locale d'assurance, ou parce que le quatrième sous-module SM4 l'a averti qu'un SLA qu'il gérait localement n'était pas respecté.

[0049] Une seconde fonction consiste à décider de l'adaptation de la configuration de l'équipement NEi, et notamment du module d'assurance embarqué MAE, en fonction des données d'informations qu'il reçoit des autres sous-modules SM1 à SM4 et des règles d'assurance qui définissent la politique locale d'assurance.

[0050] Par exemple, compte tenu des résultats d'analyse de certains paramètres du réseau, fournis par le premier sous-module SM1, le cinquième sous-module SM5 peut décider de modifier son mode d'analyse ou bien demander au deuxième sous-module SM2 d'adresser une alarme à la couche NMS de sorte qu'elle adresse de nouvelles formules d'analyse. Il peut également décider de modifier un calendrier d'analyses, par exemple effectuées par le premier sous-module SM1.

[0051] Lorsque le cinquième sous-module SM5 décide de modifier le fonctionnement (ou la configuration) de l'un des sous-modules SM1 à SM4, il lui adresse les instructions correspondantes, ce qui revient à lui demander de modifier ses calculs et/ou analyses et/ou rapports. De même, le cinquième sous-module SM5 peut, lorsque la politique locale l'y autorise, être chargé de rétroagir sur la configuration du routeur via un module de reconfiguration embarqué.

[0052] Il est important de noter que les actions de reconfiguration (ou de modification de fonctionnement) décidées par le cinquième sous-module SM5 sont définies par les règles d'assurance de la politique locale. En d'autres termes, le cinquième sous-module SM5 ne fait qu'ordonner l'exécution d'actions prédéfinies.

[0053] Le cinquième sous-module SM5 peut également appliquer des règles d'assurance choisies à des données d'information provenant d'autres équipements de réseau NEj. En effet, un équipement NEj peut être chargé de transmettre des valeurs qu'il a mesurées et/ou agrégées à un autre équipement NEi équipé d'un dispositif D chargé de traiter ces valeurs de manière à générer de nouvelles données d'information qui peuvent être ensuite remontées au NMS ou envoyées à un autre équipement NEj.

[0054] La politique d'assurance locale peut être par exemple définie comme suit :

- si la tendance, estimée par le premier sous-module SM1, du nombre de paquets abandonnés augmente de plus de 5%, alors une alarme de tendance doit être émise,
- si la bande passante utilisée, analysée par le premier sous-module SM1, demeure inférieure à un premier seuil, alors la valeur de la bande passante utilisée doit être transmise chaque jour au module de collection MCC du NMS,
- si la bande passante utilisée, analysée par le premier sous-module SM1, est comprise entre un premier et un second seuils, alors la valeur moyenne de la bande passante utilisée doit être transmise toutes les heures au module de collection MCC du NMS,
- si l'intervalle de temps, estimé par le premier sous-module SM1, pour que la bande passante utilisée atteigne le second seuil, est inférieur à deux heures, alors une alarme de tendance doit être émise,
- si la charge de calcul (CPU) de l'équipement NEi, estimée par le premier sous-module SM1, est supérieure à 80%, alors les analyses de tendances ne doivent plus être effectuées,
- si la perte de paquets, estimée par le premier sous-module SM1, augmente de 5%, alors la taille des mémoires tampons (« buffers ») doit être augmentée de 5%.

[0055] En d'autres termes la politique d'assurance locale consiste ici, d'une première part, à transmettre chaque jour la mesure de la bande passante si le comportement de l'équipement NEi est normal, mais à activer les analyses de tendance et à augmenter la fréquence d'envoi des rapports si le comportement change, d'une deuxième part, à interrompre toutes les tâches impliquant un paramètre si la valeur de ce paramètre devient critique, et d'une troisième part, à modifier le fonctionnement interne de l'équipement NEi, représenté par un paramètre, en fonction de la variation dudit paramètre.

[0056] De préférence, le module d'assurance MAE est configurable à distance, via le réseau N, par le module de configuration MC du NMS. Plus précisément, chaque sous-module SM1 à SM5 est préférentiellement configurable. Cela permet en effet d'adapter à distance, le fonctionnement de chaque sous-module SM en fonction des circonstances, par exemple du fait de l'évolution du réseau N, et/ou en fonction des besoins, par exemple lorsque le gestionnaire d'assurance centralisé

du NMS a besoin de disposer des valeurs de nouveaux paramètres, ou de paramètres agrégés complémentaires, ou encore de nouvelles analyses réelles ou prospectives (envoi de nouvelles formules ou équations).

[0057] Deux modes de configuration peuvent être envisagés. Un premier mode consiste à transmettre indirectement les données de configuration au module d'assurance MAE via une interface de langage de programmation (ou API pour « Application Programming Interface ») implantée dans l'équipement NEi, ainsi qu'éventuellement via la MIB. Cette interface API peut également, et avantageusement, permettre au module de configuration MC du NMS de configurer ou de programmer la MIB, de sorte qu'elle ne soit pas de type statique. Le module d'assurance MAE peut en effet coopérer avec la MIB pour se configurer et définir de nouvelles entrées dans la MIB, afin de pouvoir accéder à ces nouvelles entrées. Dans ce cas, on accède à la MIB via des requêtes SNMP (« GET » et « SET »), ce qui permet une configuration aisée (voire même une gestion des droits et de la sécurité des échanges avec SNMP 3) ; la MIB permet d'avoir une interface « normalisée » du fait que ses champs sont accessibles via des OIDs (pour « Object Identifiers ») bien définis. On entend ici par « Object Identifier », un identifiant objet d'une variable MIB que l'on souhaite récupérer. Un tel identifiant est généralement normalisé, par exemple par le standard RFC 1213 pour la MIB II.

[0058] Un second mode consiste à transmettre directement les données de configuration au module d'assurance MAE à l'aide de commandes dédiées, telles que des « Command Line Interface » (ou CLI).

[0059] Le module d'assurance MAE peut être réalisé sous la forme de circuits électroniques, de modules logiciels (ou informatiques), ou d'une combinaison de circuits et de logiciels.

[0060] L'invention ne se limite pas aux modes de réalisation de dispositif de gestion locale d'assurance et d'équipement de réseau décrits ci-avant, seulement à titre d'exemple, mais elle englobe toutes les variantes que pourra envisager l'homme de l'art dans le cadre des revendications ci-après.

[0061] Ainsi, dans ce qui précède on a décrit des équipements de réseau capables de gérer localement le procédé d'assurance supervisé par le NMS. Mais, on peut envisager un fonctionnement distribué dans lequel certains au moins des équipements de réseau sont équipés de dispositifs de gestion locale d'assurance simplifiés, chargés de transmettre des alarmes et/ou des données d'informations spécifiques et complémentaires à un serveur d'assurance du NMS. Bien entendu, dans ce cas les équipements de réseau demeurent capables d'adapter leur configuration en fonction des valeurs des données de gestion stockées dans leur MIB.

[0062] Par ailleurs, on a décrit un mode de réalisation dans lequel le dispositif selon l'invention était implanté ou embarqué dans l'équipement de réseau. Mais, le dispositif selon l'invention peut être agencé sous la forme d'un élément externe implanté dans un boîtier directement raccordé à l'équipement de réseau.

Revendications

1. Dispositif (D) de gestion locale d'assurance pour un équipement de réseau (NE) d'un réseau de communications (N) équipé d'un système de gestion de réseau (NMS), ledit équipement (NE) présentant une configuration choisie et comportant des moyens (MM) de mesure de valeurs de paramètre (s) du réseau et une base d'informations de gestion (MIB) propre à stocker des données de gestion représentatives desdites valeurs mesurées, **caractérisé en ce qu'**il comprend des moyens de gestion (MAE) agencés pour adapter la configuration dudit équipement (NE) en fonction d'au moins lesdites données de gestion stockées dans ladite base d'informations de gestion (MIB) et de règles, dites d'assurance, choisies, définissant une politique locale d'assurance, ladite adaptation consiste en une modification d'une politique de mesure de paramètre (s) et/ou une modification d'une politique d'envoi de rapport audit système de gestion de réseau (NMS).

2. Dispositif selon la revendication 1, **caractérisé en ce que** lesdits moyens de gestion (MAE) sont agencés pour adapter ladite configuration en fonction de données d'informations provenant d'au moins un autre équipement de réseau (NE).

3. Dispositif selon l'une des revendications 1 et 2, **caractérisé en ce que** ladite adaptation consiste en une modification de son mode de fonctionnement.

4. Dispositif selon l'une des revendications 1 à 3, **caractérisé en ce que** lesdits moyens de gestion (MAE) comprennent des moyens d'analyse (SM1) agencés pour déterminer, en fonction de certaines desdites règles d'assurance choisies, des données d'information représentatives de l'évolution temporelle, sur un intervalle choisi, de valeurs de paramètre(s) du réseau stockées dans ladite base d'informations de gestion (MIB).

5. Dispositif selon la revendication 4, **caractérisé en ce que** lesdits moyens d'analyse (SM1) sont agencés pour délivrer des données d'information représentatives d'une analyse de tendance et/ou d'une analyse de profils ou de signatures et/ou d'une analyse de discontinuité et/ou d'une agrégation de valeurs de paramètres du réseau.

6. Dispositif selon l'une des revendications 4 et 5, **caractérisé en ce que** lesdits moyens d'analyse (SM1) sont configurables.

7. Dispositif selon la revendication 6, **caractérisé en ce que** lesdits moyens d'analyse (SM1) sont agencés pour mettre en oeuvre de nouveaux calculs, relatifs aux paramètres du réseau, reçus dudit système de gestion de réseau (NMS).

8. Dispositif selon l'une des revendications 1 à 7, **caractérisé en ce que** lesdits moyens de gestion (MAE) comprennent des moyens d'alarme (SM2) propres à déclencher l'envoi d'une alarme et/ou de données d'information audit système de gestion de réseau (NMS) et/ou à au moins un autre équipement de réseau (NE), en fonction de certaines desdites règles d'assurance choisies.

9. Dispositif selon la revendication 8, **caractérisé en ce que** lesdits moyens d'alarme (SM2) sont configurables.

10. Dispositif selon l'une des revendications 8 et 9, **caractérisé en ce que** lesdites données d'information et lesdites alarmes sont représentatives d'un résultat d'analyse, effectuée par lesdits moyens d'analyse (SM1), et/ou d'une agrégation de données, effectuée par lesdits moyens d'analyse (SM1), et/ou d'une valeur de paramètre du réseau stockée dans ladite base d'informations de gestion (MIB).

11. Dispositif selon l'une des revendications 1 à 10, **caractérisé en ce que** lesdits moyens de gestion (MAE) comprennent des moyens d'observation du réseau (SM3) définissant un agent de mesure de flux de type « bout-en-bout », et agencés pour déterminer des données d'information représentatives dudit flux de type bout-en-bout en fonction de certaines desdites règles d'assurance choisies.

12. Dispositif selon la revendication 11, **caractérisé en ce que** lesdits moyens d'observation du réseau (SM3) sont configurables.

13. Dispositif selon l'une des revendications 1 à 12, **caractérisé en ce que** lesdits moyens de gestion (MAE) comprennent des moyens de gestion d'accords de niveau de service (SM4) agencés pour déterminer des données d'information représentatives de ladite gestion des accords en fonction de certaines desdites règles d'assurance choisies.

14. Dispositif selon la revendication 13, **caractérisé en ce que** lesdits moyens de gestion d'accords de niveau de service (SM4) sont configurables.

15. Dispositif selon l'une des revendications 2 à 14, **caractérisé en ce que** lesdits moyens de gestion (MAE) comprennent des moyens de contrôle (SM5) propres à gérer le fonctionnement desdits moyens d'analyse (SM1), desdits moyens d'alarme (SM2), desdits moyens d'observation du réseau (SM3) et desdits moyens de gestion d'accords de niveau de service (SM4), en fonction de certaines au moins desdites règles d'assurance choisies.

16. Dispositif selon la revendication 15, **caractérisé en ce que** lesdits moyens de contrôle (SM5) sont alimentés en données d'information par lesdits moyens d'analyse (SM1) et/ou lesdits moyens d'observation du réseau (SM3) et/ou lesdits moyens de gestion d'accords de niveau de service (SM4), et sont agencés pour ordonner auxdits moyens d'alarme (SM2) de générer des alarmes et/ou des rapports en cas de détection du non respect d'une règle d'assurance par des données d'information reçues.

17. Dispositif selon l'une des revendications 15 et 16, **caractérisé en ce que** lesdits moyens de contrôle (SM5) sont agencés sous la forme d'un moteur de règles (« rule engine ») stockant lesdites règles d'assurance choisies.

18. Dispositif selon l'une des revendications 15 à 17, **caractérisé en ce que** lesdits moyens de contrôle (SM5) sont configurables.

19. Dispositif selon l'une des revendications 1 à 18, **caractérisé en ce que** lesdits moyens de gestion (MAE) sont propres à être configurés par ledit système de gestion de réseau (NMS) via une interface de langage de programmation (API) dudit équipement (NE).

20. Dispositif selon l'une des revendications 1 à 19, **caractérisé en ce que** lesdits moyens de gestion (MAE) sont propres à être configurés par ledit système de gestion de réseau (NMS) via une interface de langage de programmation (API) dudit équipement (NE) et via ladite base d'informations de gestion (MIB).

21. Dispositif selon l'une des revendications 19 et 20, **caractérisé en ce que** lesdits moyens d'analyse (SM1) et/ou lesdits moyens d'alarme (SM2) et/ou lesdits moyens d'observation du réseau (SM3) et/ou lesdits moyens de contrôle (SM5) et/ou lesdits moyens de gestion d'accords de niveau de service (SM4) sont propres à être configurés par ledit système de gestion de réseau (NMS), via ladite interface de langage de programmation (API).

22. Dispositif selon l'une des revendications 19 et 20, **caractérisé en ce que** lesdits moyens d'analyse (SM1) et/ou lesdits moyens d'alarme (SM2) et/ou lesdits moyens d'observation du réseau (SM3) et/ou lesdits moyens de contrôle (SM5) et/ou lesdits moyens de gestion d'accords de niveau de service (SM4) sont propres à être configurés par ledit sys-

tème de gestion de réseau (NMS), via ladite interface de langage de programmation (API) et via ladite base d'informations de gestion (MIB).

23. Dispositif selon l'une des revendications 1 à 22, **caractérisé en ce que** lesdits moyens de gestion (MAE) sont propres à être configurés par ledit système de gestion de réseau (NMS) via des commandes dédiées.

24. Dispositif selon la revendication 23, **caractérisé en ce que** lesdits moyens d'analyse (SM1) et/ou lesdits moyens d'alarme (SM2) et/ou lesdits moyens d'observation du réseau (SM3) et/ou desdits moyens de gestion d'accords de niveau de service (SM4) et/ou lesdits moyens de contrôle (SM5) sont agencés pour être configurés par ledit système de gestion de réseau (NMS) via des commandes dédiées.

25. Dispositif selon l'une des revendications 23 et 24, **caractérisé en ce que** lesdites commandes sont de type « Command Line Interface ».

26. Equipement de réseau (NE) pour un réseau de communications (N) équipé d'un système de gestion de réseau (NMS), ledit équipement (NE) présentant une configuration choisie et comportant des moyens (MM) de mesure de valeurs de paramètre (s) du réseau et une base d'informations de gestion (MIB) propre à stocker des données de gestion représentatives desdites valeurs mesurées, **caractérisé en ce qu'**il comprend un dispositif (D) selon l'une des revendications précédentes.

27. Equipement selon la revendication 26, **caractérisé en ce qu'**il comprend une interface de langage de programmation (API), et **en ce que** ladite base d'informations de gestion (MIB) est propre à être configurée par ledit système de gestion de réseau (NMS) via ladite interface de langage de programmation (API).

28. Equipement selon l'une des revendications 26 et 27, **caractérisé en ce qu'**il comprend une interface de langage de programmation (API), et **en ce que** ladite base d'informations de gestion (MIB) est propre à être programmée par ledit système de gestion de réseau (NMS) via ladite interface de langage de programmation (API).

29. Equipement selon l'une des revendications 26 à 28, **caractérisé en ce qu'**il est choisi dans un groupe comprenant au moins les routeurs, les commutateurs et les pare-feu.

30. Réseau de communications (N) comportant un système de gestion de réseau (NMS), **caractérisé en ce qu'**il comprend une multiplicité d'équipements de réseau (NE) selon l'une des revendications 26 à 29.

31. Réseau selon la revendication 30, **caractérisé en ce que** chaque équipement (NE) est agencé pour délivrer audit système de gestion de réseau (NMS) des alarmes et/ou des données d'information de types différents.

32. Utilisation des dispositif, équipement et réseau de communications selon l'une des revendications précédentes dans les technologies réseaux devant être gérées.

33. Utilisation selon la revendication 32, **caractérisé en ce que** lesdites technologies réseaux sont choisies dans un groupe comprenant les réseaux de transmission, en particulier de type WDM, SONET et SDH, de données, en particulier de type Internet-IP et ATM, et de voix, en particulier de type classique, mobile et NGN.

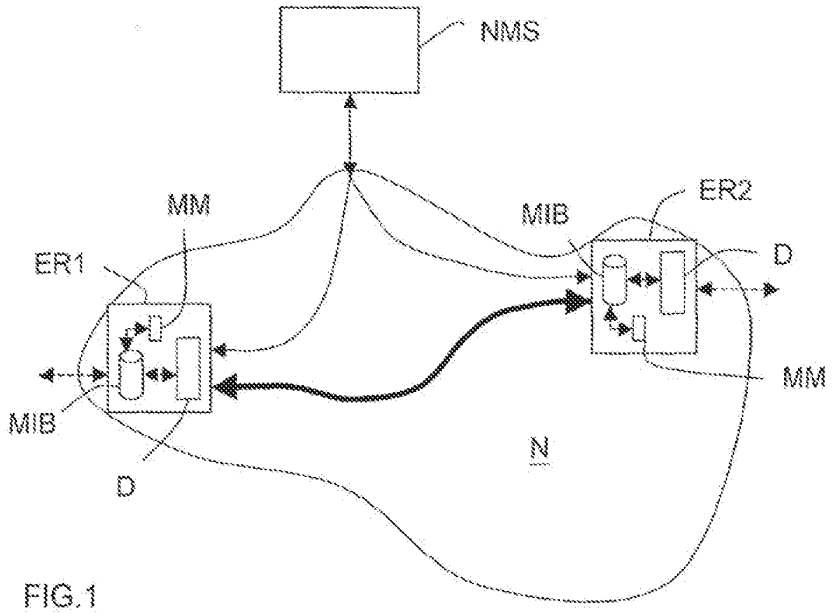

NMS
MM
MIB
ER2
ER1
D
MM
MIB
N
D

FIG.1

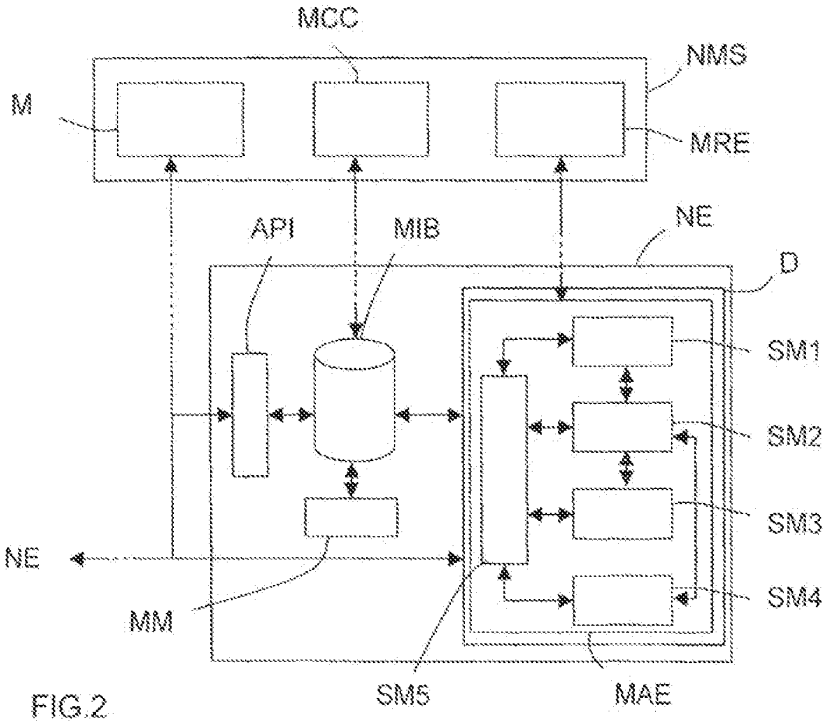

MCC
NMS
M
MRE
API
MIB
NE
D
SM1
SM2
SM3
SM4
NE
MM
SM5
MAE

FIG.2

Office européen des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numéro de la demande
EP 04 29 0706

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.7)
X	HO L L ET AL: "REAL-TIME PERFORMANCE MONITORING AND ANOMALY DETECTION IN THE INTERNET: AN ADAPTIVE, OBJECTIVE-DRIVEN, MIX-AND-MATCH APPROACH" BELL LABS TECHNICAL JOURNAL, WILEY, CA, US, vol. 4, no. 4, octobre 1999 (1999-10), pages 23-40, XP000902469 ISSN: 1089-7089 * abrégé * * page 24, colonne de droite - page 35, colonne de droite * * page 38, colonne de droite, alinéa 2 - page 39, colonne de gauche, alinéa 1 *	1-33	H04L12/24 H04L12/26
X	US 6 272 537 B1 (KEKIC MIODRAG M ET AL) 7 août 2001 (2001-08-07) * colonne 7, ligne 20 - ligne 65 * * colonne 8, ligne 42 - colonne 9, ligne 8 * * colonne 13, ligne 50 - colonne 85, ligne 11 *	1-33	DOMAINES TECHNIQUES RECHERCHES (Int.Cl.7) H04L
X	FR 2 774 191 A (INST EURECOM) 30 juillet 1999 (1999-07-30) * abrégé * * page 1, ligne 2 - ligne 8 * * page 3, ligne 5 - ligne 7 * * page 3, ligne 12 - ligne 17 * * page 4, ligne 12 - ligne 17 * * page 5, ligne 1 - page 6, ligne 3 * * page 7, ligne 1 - ligne 12 * * page 11, ligne 12 - page 16, ligne 6 * * figures 2,6,7 *	1-33	

-/--

Le présent rapport a été établi pour toutes les revendications

Lieu de la recherche	Date d'achèvement de la recherche	Examinateur
Munich	27 juillet 2004	Le Bras, P

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : arrière-plan technologique
O : divulgation non-écrite
P : document intercalaire

T : théorie ou principe à la base de l'invention
E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date
D : cité dans la demande
L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

EPO FORM 1503 03.82 (P04C02)

Office européen des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numéro de la demande
EP 04 29 0706

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.7)
X	US 6 085 243 A (BANTHIA PRAKASH C ET AL) 4 juillet 2000 (2000-07-04) * abrégé * * colonne 1, ligne 17 - ligne 20 * * colonne 3, ligne 52 - colonne 5, ligne 10 * * colonne 5, ligne 52 - ligne 54 * * colonne 5, ligne 65 - colonne 6, ligne 26 * * colonne 6, ligne 44 - ligne 46 * * colonne 7, ligne 12 - ligne 23 * * colonne 8, ligne 14 - colonne 10, ligne 58 * * colonne 12, ligne 2 - ligne 4 * * colonne 14, ligne 11 - ligne 21 *	1-33	

DOMAINES TECHNIQUES RECHERCHES (Int.Cl.7)

Le présent rapport a été établi pour toutes les revendications

Lieu de la recherche	Date d'achèvement de la recherche	Examinateur
Munich	27 juillet 2004	Le Bras, P

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : arrière-plan technologique
O : divulgation non-écrite
P : document intercalaire

T : théorie ou principe à la base de l'invention
E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date
D : cité dans la demande
L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

EPO FORM 1503 03.82 (P04C02)

ANNEXE AU RAPPORT DE RECHERCHE EUROPEENNE RELATIF A LA DEMANDE DE BREVET EUROPEEN NO. EP 04 29 0706

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche européenne visé ci-dessus.
Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets.

27-07-2004

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 6272537 B1	07-08-2001	AUCUN	
FR 2774191 A	30-07-1999	FR 2774191 A1	30-07-1999
US 6085243 A	04-07-2000	AU 6601198 A	03-07-1998
		CA 2275235 A1	18-06-1998
		EP 0956680 A1	17-11-1999
		GB 2335124 A	08-09-1999
		WO 9826541 A1	18-06-1998
		US 5922044 A	13-07-1999
		US 6069274 A	28-12-1999
		US 6108782 A	22-08-2000
		WO 9843170 A1	01-10-1998

EPO FORM P0459

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No 12/82

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
10 July 2008 (10.07.2008)

(10) International Publication Number
WO 2008/080430 A1

(51) International Patent Classification:
H04L 29/06 (2006.01)

(21) International Application Number:
PCT/EP2006/012006

(22) International Filing Date:
29 December 2006 (29.12.2006)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **TELECOM ITALIA S.P.A.** [IT/IT]; Piazza degli Affari, 2, I-20123 Milano (IT).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **COSTA, Luciana** [IT/IT]; Telecom Italia S.p.A, Via G. Reiss Romoli, 274, I-10148 Torino (IT). **FREGUGLIA, Giorgio** [IT/IT]; Telecom Italia S.p.A., Via G. Reiss Romoli, 274, I-10148 Torino (IT). **FROSALI, Federico** [IT/IT]; Telecom Italia S.p.A., Via G. Reiss Romoli, 274, I-10148 Torino (IT).

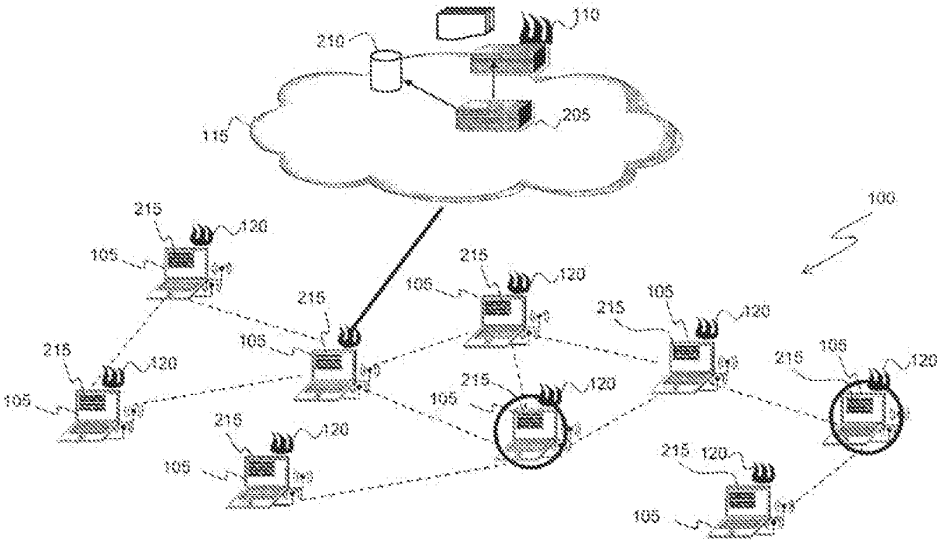
(74) Agent: **BATTIPEDE, Francesco**; Telecom Italia S.p.A., Piazza L. Einaudi, 8, I-20124 Milano (IT).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) Title: METHOD AND SYSTEM FOR ENFORCING SECURITY POLICIES IN MANETs



(57) Abstract: A method of enforcing security policies in a mobile ad-hoc network, comprising: - entrusting at least one first network node, along a data traffic route from a data traffic origin node to a data traffic destination node, with the enforcing of predefined security policies on said data traffic; - entrusting at least one second network node, distinct from said first network node, with the control of the enforcement of the security policies by said first network node.

WO 2008/080430 A1

METHOD AND SYSTEM FOR ENFORCING SECURITY POLICIES IN MANETs

DESCRIPTION

Background of the invention

Field of the invention

The present invention generally relates to the field of telecommunications, and particularly to Mobile Ad-hoc NETworks (MANETs). Even more particularly, the invention relates to the implementation and enforcing of security policies in MANETs. Specifically, the invention concerns a method and a system adapted to detect and react to situations wherein the enforcement of security policies set forth for a MANET is missing or improper.

Description of the related art

The heterogeneous nature and the mobility characteristics of a MANET makes it particularly difficult to control the behavior of the nodes that make up the network.

Nowadays, the acceptance of one or more network security rules by a generic MANET node as a condition for allowing the node accessing and becoming part of the network does not prevent that node from violating said rules later, and thus escaping from the control upon it. A node that, after an authentication procedure, has been granted access to an *ad-hoc* network is potentially able to perform any action, such as stop properly routing traffic received from and directed to other MANET nodes, allowing access to services to which some node is not authorized, allowing access to other nodes by revealing the necessary security information, like the ciphering keys.

For these reasons, distributed security mechanisms need to be defined for MANETs, aiming at making it more difficult to corrupt the network nodes.

The problem of the enforcement of security policies in *ad-hoc* networks has already been addressed.

For example, in the thesis work "Adaptive distributed firewall using intrusion detection" of L. Strand of the Oslo University, reference is made to a distributed firewalling architecture in which a master (a management station) distributes the security policies to clients, which are the nodes making up the network; the clients apply the security policies and notify the master of the results of the security policies application.

The described architecture is not expressly conceived for MANETs, but the latter is one of the environments considered as a possible application. The interaction with a distributed intrusion detection system is also considered, which, in case of necessity, allows to command a reconfiguration of the firewall of a certain node.

In J. Grant *et al.* "Distributed firewall technology", Signal 9 Solutions Inc., 2000, a distributed firewalling architecture is considered wherein a Distributed Firewall Server (DFS) distributes the security policies, valid for a certain network, to all the clients belonging thereto (possibly in a personalized way),

Each client, in normal operation, creates locally log files of its operation (also related to possible malfunctioning of the firewall) and of the traffic passing through that node; the logged data are then sent to the DFS, that correlates them for detecting possible inconsistencies with respect to the security policies deployed. The DFS, if possible, may also attempt to sniff the traffic, to check the proper implementation of the security policies deployed.

A different approach is described in S. Rajavaram *et al.*, "Neighboring watch: an intrusion detection and response protocol for mobile ad-hoc networks", University of Maryland, October 2002, where an architecture for detecting intrusions and anomalous behavior of the nodes belonging to a MANET is discussed. The possibility is exploited that a client (controller) being in the neighbourhood of other nodes "listens", over the wireless link, to the communications of said neighbouring nodes, so as to determine the presence of undesired or non-compliant traffic. The detection is made individually by every client, and the (possible) countermeasures may be applied locally, in a passive way, by exclusion of the malicious node from the announcements performed by the controller with the routing protocol, or in active way, notifying the violation to a master-cluster station that sets-up an election system involving all the MANET nodes for deciding together whether or not to exclude a certain client.

Summary of the invention

The Applicant has observed that the adoption of a distributed security mechanism is not, *per-se*, sufficient to ensure that each network node behaves properly. Moreover, from the viewpoint of a network operator/administrator, the adoption of a totally distributed security mechanism may be unacceptable, because incompatible with the network control capability that an administrator may want to have.

In the work of L. Strand, there is no real and reliable way of controlling the application of the security policies. From one hand, it is stated that it is the network node applying the policies that has to notify to the management station whether the policy application was successful: regretfully, if that node is corrupted, or pursues different plans, it may send to the management station false information. From the other hand, even if a distributed intrusion detection mechanism is envisaged, which might affect the firewall reconfiguration on the nodes, nothing is said about how this could be done; moreover, it seems that the proposed way is to employ the intrusion detection system to reveal classic intrusions, instead of controlling the proper application of the firewalling rules.

In the work of J. Grant *et al.*, which is not expressly conceived for MANETs, no efficient method of control of the security policies applied by the different clients is provided. The proposed solution relies in part on the communications between the clients and the DFS, but the reported data might be totally false. Also in respect of the envisaged possibility that the DFS directly sniffs the traffic for controlling the proper application of the policies, it is not explained how this could be done; moreover, within a MANET it is practically impossible that a single station, even if located within the MANET, can intercept all the network traffic, because in a wireless network any station may at most listen to the traffic passing within its radio

range.

In the work of S. Rajavaram *et al.*, the detection of anomalous behaviors is not focused on the control of the implementation of firewalling policies, rather on the generic control of a node. No mechanism is described for the selection of the nodes to be controlled by a controller node: this means that a node controls all the clients within its radio range, without limitations, but this might cause a great waste of the controller node's resources. Additionally, the proposed architecture does not envisage the presence of a real management node that tracks the overall behavior of the network.

The Applicant believes that there is the need of providing, for MANETs managed by a network operator/administrator, security mechanisms capable of exercising a control on the behavior of the network nodes, and also capable of allowing to check the proper application of the security policies intended to be implemented.

The Applicant has found that this need can be satisfied by means of a security mechanism which is based, from one side, on the centralization of the control mechanisms, through which the network is coordinated and managed, and, from another side, on the distribution of the security functions across the different network nodes.

In particular, the security mechanism according to an embodiment of the present invention allows to control the proper application, by every network node, of the security policies set forth by a network administrator, and to identify possible violations. This is achieved by entrusting the nodes of the network with the task of controlling the enforcement of the security policies, in order to detect cases of violations/improper enforcement of the security policies by other network nodes, and to identify the network nodes that are responsible of the security policies misapplication.

By distributing the control function to several, potentially all the network nodes, the data processing burden on the generic node is reduced, and the performance is thus improved, avoiding at the same time the formation of malicious "agreements" between network nodes, which could make the control vain; the overall network security is thus increased.

Essentially, the present invention proposes a mechanism for identifying the network nodes that are eligible for being controlled: advantageously, these nodes are a subset of the nodes of the network, and particularly they are the nodes that, due to the network topology, have the task of routing traffic from/to other network nodes.

In turn, the nodes that will have to perform the task of controllers are identified by an interaction with the routing protocol, designed to reduce, as far as possible, the number of nodes to be controlled by each controller node. Indeed, to verify the proper application of the security policies defined for the MANET, particularly to detect unauthorized traffic, each node should monitor all the traffic involving the nodes that are within its radio range, but this would imply a substantial waste of resources (particularly energy, which in mobile, wireless nodes is always precious). According to an embodiment of the present

invention, the number of controls to be carried out is significantly limited, by properly choosing the nodes to be controlled, to the extent that, at best, a node may continuously control just a single other node. Advantageously, a generic node may in general be controlled by more than one controller nodes (depending on the network topology).

The controller nodes are able to identify which node(s) has not properly enforced the security policies, as well as which node(s) violated the security policies, exploiting the missing enforcement of the security policies by network nodes that are entrusted with the enforcement thereof on the violator node.

The invention also provides for countermeasures to be taken once one or more nodes have been identified that violated or did not enforce the security policies.

Further, according to an embodiment of the present invention, selfish network nodes may be identified (and contrasting actions can accordingly be taken), *i.e.* nodes that tend to participate to the MANET in a utilitarian way, propagating only their own data packets, and not forwarding (discarding) data packets originated by, and addressed to other nodes.

According to an aspect of the present invention, a method of enforcing security policies in a mobile *ad-hoc* network is provided, comprising:

- entrusting at least one first network node, along a data traffic route from a data traffic origin node to a data traffic destination node, with the enforcing of predefined security policies on said data traffic;

- entrusting at least one second network node, distinct from said first network node, with the control of the enforcement of the security policies by said first network node.

According to another aspect of the present invention, a system for enforcing security policies in a mobile *ad-hoc* network, comprising:

- at least one first network node, along a data traffic route from a data traffic origin node to a data traffic destination node, entrusted with the enforcing of predefined security policies on said data traffic;

- at least one second network node, distinct from said first network node, entrusted with the control of the enforcement of the security policies by said first network node.

Brief description of the drawings

The features and advantages of the present invention will result apparent by reading the following detailed description of an embodiment thereof, provided merely by way of non-limitative example, and referring to the annexed drawings, wherein:

Figure 1 pictorially shows a scenario of a MANET where the present invention can be applied;

Figure 2 schematically shows network elements implementing a Distributed Firewall Control (DFC) architecture, according to an embodiment of the present invention;

Figure 3 describes a format of HELLO messages exploited by a routing protocol adopted in the MANET;

Figure 4 schematically shows an example of application of a DFC mechanism implemented by the DFC architecture of Figure 2;

Figure 5A shows a format of Topology Control (TC) messages used by the routing protocol;

Figure 5B schematically shows the information that a DFC client of the DFC architecture maintains in order to detect the nodes whose behavior is not compliant to MANET security policies;

Figure 5C schematically shows the information included in a routing table kept by a DFC client;

Figure 6 schematically shows a DFC selection mechanism in case of a loop;

Figure 7 is a schematic flowchart of a node identification mechanism for identifying the network node that has not enforced the prescribed security policies, according to an embodiment of the present invention;

Figures 8A to 8F schematically show examples of the identification of nodes that have not enforced the security policies;

Figure 9 is a schematic flowchart of a node quarantine mechanism according to an embodiment of the present invention;

Figures 10A and 10B are schematic flowcharts of a process of management of alarm messages used in the DFC architecture, according to an embodiment of the present invention;

Figure 11 schematically shows a content of a notification alarm message used in an embodiment of the present invention;

Figure 12 schematically shows a content of a quarantine alarm message used in an embodiment of the present invention;

Figure 13 schematically shows a content of an alert alarm message used in an embodiment of the present invention;

Figures 14A and 14B show a content of global repair request and accept alarm messages used in an embodiment of the present invention;

Figure 15 describes an architecture of a selfish node detection mechanism according to an embodiment of the present invention; and

Figure 16 shows a content of a selfish node quarantine message used in an embodiment of the present invention;

Detailed description of the preferred embodiment of the invention

In Figure 1, a scenario in which the present invention is advantageously applied is schematically shown. In the drawing, reference numeral 100 denotes a MANET comprising a plurality of mobile network nodes 105, in wireless communications relationship, where the generic network node 105 communicates with other network nodes through neighbouring network nodes 105 that, at a certain moment, fall within its radio range. In the drawing, two nodes that are in direct connection with one another are depicted as connected by a dashed line. Since in general two network nodes cannot in direct connection, because

they cannot fall within their respective radio ranges, each node **105** of the MANET acts not only as a traffic originator and/or a traffic destination, but also as a traffic router, routing traffic from/to other nodes of the network. Thus, even if the radio range of the network nodes is relatively limited, the network can cover a wide area, because nodes that are not directly interconnected can nevertheless communicate with each other *via* a multi-hop transmission.

Reference numeral **110** denotes a functional entity, like a server, connected to the MANET **100** through an infrastructured network **115**, and configured to centrally manage the deployment to, and implementation of security policies in the MANET nodes **105**. The infrastructured network may include one or more Public Switched Telephone Network (PSTN) and/or one or more mobile telephony networks like GPRS or UMTS networks; the nature and type of the infrastructured network **115** is *per-se* not limitative to the present invention. The MANET **100** may in particular constitute an "extension" of the infrastructured network **115**, by means of which services offered by the network operator to users of the infrastructured network **115** can also be offered to users of the mobile network nodes **105**. The MANET **100** may also operate as a replacement of the infrastructured network **115** in case of failure of the latter.

In particular, the server **110** may be managed by an administrator of the MANET **100**, for example an administrator operating on behalf of the operator of the infrastructured network **115** MANET **100**. By means of the server **110**, the network administrator keeps a control on the traffic exchanged through the MANET **100**, defining security policies to be applied by the MANET nodes **105**; the security policies may for example be deployed dynamically *via* the infrastructured network **115** and the MANET **100**.

According to the present invention, it is ensured that the desired security policies are effectively applied and respected, *i.e.* that the security policies are enforced.

According to an embodiment of the present invention, a Distributed Firewalling Architecture (DFA) is implemented in the MANET **100**, adapted to set-up a distributed firewalling mechanism, according to which the firewalling operations are entrusted to the MANET nodes **105**. The generic MANET node **105** behaves as a client (hereinafter referred to as a Distributed Firewalling Architecture client or, shortly, "DFA client") in respect of the server **110** (hereinafter referred to as the "DFA server"). The client resident in the generic MANET node **105**, schematically depicted in the drawing and denoted **120**, is adapted to dynamically receive from the DFA server **110** security policies to be applied in the MANET, to interpret the received security policies and to translate them into lower-level traffic filtering rules (*e.g.*, Internet Protocol – IP - filtering rules), used for filtering the traffic that is routed by the MANET nodes, as will be described in detail later.

According to an embodiment of the present invention, the generic MANET node **105** does not filter its own traffic, *i.e.* the traffic it originates or that is destined thereto, rather it filters the traffic of the other MANET nodes, particularly of the neighbouring MANET nodes; in other words, considering the

generic MANET node **105**, its traffic is filtered by one or more of the neighbouring MANET nodes.

In particular, according to an embodiment of the present invention, an efficient traffic filtering policy is implemented, which provides for entrusting the network traffic filtering operation to a subset of MANET nodes, particularly to the neighbouring nodes of a node that is either the source of the data traffic or the destination of the data traffic. In this way, it is avoided that every MANET node has to apply the security policies, *i.e.* the data traffic filtering rules valid for all the MANET nodes, and has to compare to such rules every data packet received before routing it. In other words, according to an embodiment of the present invention, the filtering of a data packet sent by a certain source or origin node, like the network node denoted **105s** in **Figure 1**, towards a certain destination, like the network node **105d** in the drawing, is carried out only by the (DFA client resident on the) node **105a** that, along the route **125** followed by the data packet, immediately follows the source node, and by the node **105b** that, along said route, immediately precedes the destination node.

In particular, in an embodiment of the present invention, indications about the network topology are exploited to identify which MANET nodes have to filter the data traffic; in particular, said network topology indications are included in and propagated by routing messages from which the generic DFA client **120** can derive, in real time, which are the MANET nodes with which it can communicate. In particular, the knowledge of the distance from each node, in terms of number of hops, enables the DFA client **120** resident on the generic MANET node **105** to classify the nodes of the MANET as neighbouring nodes and distant nodes; for the purposes of the present description and claims, by "neighbouring node" it is meant a MANET node that is one hop distant from the generic node, whereas by "distant node" it is meant a MANET node more than one hop distant from the generic node. In this way, the DFA client **120** resident on the generic MANET node **105** is adapted to dynamically configure itself, adding and/or removing security policy rules based on the changes in the network topology, and actuating only the rules that are associated with the nodes it knows, *i.e.* its neighbouring nodes.

Entrusting the enforcement of the security policies defined in respect of a certain MANET node **105**, like for example the node **105s** in **Figure 1**, to the (DFA client **120** resident on the) neighbouring nodes, *i.e.* the node **105a** in the example of **Figure 1**, avoids that the remaining, successive nodes on the route **125** followed by a data packet sent by the node **105s** (like the node **105c** in the shown example) have to perform filtering operations on the same data traffic, because this filtering operation is useless, since assumed that the filtering has been made by the node **105a** the traffic inherently complies with the prescribed security policies.

The further filtering entrusted to the node that, like the node **105b** in the example of **Figure 1**, immediately precedes the data packet destination node **105d** is not essential, but it increases the overall network security.

In addition to the distributed firewalling architecture and mechanism described above, according

to an embodiment of the present invention, a distributed firewall control architecture is implemented in the MANET **100**, adapted to set-up a distributed control mechanism for detecting the presence, in the MANET **100**, of traffic not compliant with a certain security policy, and, in case such non-compliant traffic is detected, to identify the node(s) which has not properly enforced the security policy, as well as to identify the node(s) that violated the security policy, and to consequently take proper countermeasures. The distributed control mechanism allows avoiding that unauthorized traffic, generated by a node that violated the security policies, and not properly filtered by the client DFA on the node entrusted with the security policies enforcement (the node that immediately follows the source node), is routed through several network nodes before being filtered out, for example by the node that immediately precedes the destination node, with a consequent saving of network nodes' power and bandwidth resources.

According to an embodiment of the present invention, as schematically shown in **Figure 2**, the distributed firewall control architecture comprises a distributed firewall control server (hereinafter shortly referred to as "DFC server") **205**, possibly, albeit not limitatively, co-located with the DFA server **110**. The DFC server **205** is configured to centrally manage and coordinate a distributed firewall control mechanism, taking proper countermeasures in respect of MANET nodes that do not properly enforce the security policies and nodes that violate the security policies defined for the MANET; the DFC server **205** interacts with the DFA server **110** and, possibly, with further entities forming a security platform of the MANET, like for example one or more authentication servers **210**.

Like in the scenario of **Figure 1**, each MANET node **105** acts as a client (DFA client) of the DFA server **110**, *i.e.* as an enforcer of the firewalling policies defined for the nodes that, at a certain time, are its neighbours. In addition to the DFA client, a DFC client **215** is resident on each MANET node **105**, adapted to detect the proper application of the security policies by the DFA clients, *i.e.* by the security policies enforcers.

Additionally, on the generic MANET node **105** one or more applications designed to implement the desired routing protocol are assumed to be resident (not shown in the drawing), adapted to set up the data packets routes for reaching nodes that are distant more than one hop. In particular, said routing protocol is assumed to be a proactive routing protocol. As known to those skilled in the art, routing protocols defined for MANETs are generally classified as proactive protocols and reactive protocols, depending on how the routes towards the destination nodes are determined; in proactive protocols, every node periodically propagates to the network its topological information so that each node in the MANET is able to build the network map and to calculate a path to reach every node. With reactive protocols, the generic node calculates a path towards another node only at the time the former node has to send data to the latter node.

In particular, according to an embodiment of the present invention, the routing protocol adopted is the Optimized Link State Routing (OLSR) protocol, described for example in T. Clausen, P. Jacquet,

"Optimized Link State Routing Protocol (OLSR)", RFC 3626, October 2003. The OLSR protocol uses a broadcast algorithm exploiting a subset of the MANET nodes **105**, referred to as the MultiPoint Relay (MPR), specifically for the propagation of the routing messages. Each node that implements the OLSR routing protocol builds its own set of MPR nodes (the "MPR set"), by means of a procedure called "neighbour detection", and gets indications about the lists of nodes that are distant one or two hops therefrom; then, the considered node elects as MPR node one of the neighbouring nodes with which there is a bidirectional link (a "symmetric neighbour"), which allows the considered node to reach another node distant two hops. In this way, the union of the areas covered by the radio ranges of all the MPR nodes elected by one generic node covers all the nodes that are distant two hops therefrom. Each node sends to its neighbouring nodes the respective set of elected MPR nodes through particular OLSR protocol messages, referred to as "*HELLO*" messages; the structure of a *HELLO* message is pictorially shown in **Figure 3**. The *HELLO* message, globally denoted **300**, includes one or more fields **305** called "*Neighbour Interface Address*", wherein each node inserts the network addresses (IP addresses) of its neighbouring nodes, distant one hop, and a field **310** called "*Neighbour Type*", wherein each node defines if the node with neighbour interface address is an MPR for it . For example, referring to scenario shown in **Figure 4**, wherein an exemplary network topology **400** is depicted, node **405a** inserts into the fields **305** of the *HELLO* messages it sends the addresses of nodes **405b**, **405c**, **405d** and **405e**; among these nodes, node **405a** elects as MPR node only the node **405e**, thus, in the *HELLO* messages, it sets for this node the field **310** to a value "*MPR_NEIGH*", to indicate that the node **405e** is an MPR.

The nodes that are elected MPR nodes, in addition to sending their own *HELLO* messages, periodically send "*TOPOLOGY CONTROL*" ("*TC*") messages; the structure of a *TC* message is pictorially shown in **Figure 5A**: the message, globally denoted **500**, includes a plurality of fields **505** called "*Advertised Neighbour Main Address*", each one adapted to contain the address of one network node that has elected the considered node as its MPR. In this way, each network node can build its own routing table; the routes towards all the possible destinations are calculated passing through the MPR nodes elected by each node.

According to an embodiment of the present invention, the firewall control mechanism implemented by the DFC architecture advantageously exploits the principles of operation of the OLSR routing protocol and the correspondingly obtained network topological indications, in order to limit the number of controls to be performed to a subset of the MANET nodes. In particular, also the MANET data traffic is caused to pass through the nodes that are elected MPR nodes in respect of the OLSR protocol. This is achieved by causing every node, when building the respective routing table, to configure as next-hop for reaching a desired destination one of the MPR nodes which have been elected by that node. In this way, a first optimization of the DFC mechanism is achieved, that allows performing the firewall control function only on the MPR nodes, *i.e.* only on those nodes that actually handle the routing of data traffic

from/to the other nodes. The MPR nodes form altogether a relatively small subset of all the MANET nodes (in **Figure 4**, the MPR nodes are shown as encircled, and the subset of MPR nodes is denoted **450**). Thus, the subset of MPR nodes becomes the set of the DFA clients **120** that are subjected to controls in order to guarantee that the security policies defined for the MANET are actually applied and enforced.

Additionally, the DFC client **215** resident on the generic network node is adapted to choose, within the above-mentioned subset **450** of DFA clients **120** to be controlled, that or those DFA clients on which to perform the control, by means of a nodes selection mechanism.

According to an embodiment of the present invention, the nodes selection mechanism enables the generic DFC client **215** to:

- limit the number of controls it has to perform to a small, possibly the smallest number of DFA clients **120** (ideally, a DFC client may have to control just one DFA client); and

- guarantee that the choice of the node to control does not jeopardize the possibility of controlling a different DFA client **120**; in other words, all the DFA clients belonging to the above-mentioned subset **450** of the MPR nodes should be controlled by at least one DFC client.

Hereinafter, the DFC mechanism according to an embodiment of the present invention is described in detail, applied to the exemplary MANET shown in **Figure 4**. It is assumed that the nodes denoted **405a**, **405e**, **405f**, **405g** and **405h** (shown as encircled) are the elected MPR nodes elected by the routing process of the OLSR; the nodes **405a**, **405e**, **405f**, **405g** and **405h** thus represent the subset **450** of the DFA clients **120** that are to be controlled. Each DFC client **215** knows this subset of nodes **405a**, **405e**, **405f**, **405g** and **405h** based on the topological indications acquired by the routing protocol.

Essentially, the DFC mechanism includes three phases.

In a first phase, based on the subset of nodes thus built, each DFC client **215** chooses one DFA client **120** (*i.e.*, one node) within the subset of nodes to be controlled, using the above mentioned selection mechanism.

A second phase and a third phase, described in detail later on, concern actions to be taken in case it is ascertained that the security policies are not properly implemented; the actions that, according to an embodiment of the present invention, can be taken are a local repair action and a global repair action.

In an embodiment of the present invention, the selection mechanism adopted for selecting the DFA clients **120** to be controlled is the following.

Each DFC client **215**, residing on a certain network node **105**, firstly assesses the number of MPR nodes that have been elected by the network node where such DFC client **215** resides.

If the number of elected MPR nodes is equal to 1, then the DFA client to be controlled is the sole MPR node that has been elected. Referring to the example of **Figure 4**, the node **405c** is assumed to have elected as MPR only the node **405a**, thus the DFC client **215** resident on node **405c** controls the DFA client **120** resident on node **405a**.

If instead the number of elected MPR nodes is greater than 1, the DFC client **215** assesses whether the node where it resides is involved in a loop, *i.e.* it is part of a closed path, originating at and terminating in a same node; in the example depicted in **Figure 6**, a loop **600** exists, originating at node **605a**, passing through nodes **605b**, **605c**, **605d** and **605e**, and terminating in node **605a**. The assessment of the existence of loops by the generic DFC client is for example carried out using the MPR selector set transmitted by every MPR node in the *TC* messages it sends (the MPR selector set is the group of node addresses contained in the fields *Advertised Neighbour Main Address* **505** of the *TC* message **500** depicted in **Figure 5A**); as mentioned in the foregoing, this information allows building up a path from a certain source node to a certain destination node.

If the DFC client ascertains that there is a loop originating at and terminating in the node where it resides, like in the case of the DFC client **215** residing on node **605a** of **Figure 6**, then the DFC client **215** decrees that the node where it resides is involved in a loop.

In case no loops are identified, the DFC client **215** chooses to control, among the MPR nodes elected by the node where it resides, the MPR node through which the lowest number of nodes distant two hops can be reached. Referring to the example of **Figure 4**, let it be assumed that the node **405e** has elected as MPR the nodes **405a** and **405f**; through the node **405a**, the node **405e** can reach the nodes **405b**, **405c** and **405d**, distant from it two hops, whereas through the node **405f** it can only reach the node **405g**: thus, the DFC client resident on the node **405e** chooses to control the DFA client **120** resident on the node **405f**. If two or more MPR nodes allow reaching the same number of nodes distant two hops, then the DFC client **215** chooses as DFA client to control the one resident on a node which, in turn, has elected the lowest number of MPR nodes. If even in this way two or more nodes are identified, then the choice is made randomly. This is for example the case of node **405f** of **Figure 4**, which has elected as MPR nodes the nodes **405e** and **405g**; through the node **405e**, the node **405f** can reach the nodes **405a**, **405i** and **405m**, whereas through the node **405g** it can reach the nodes **405h**, **405n** and **405p**, all distant two hops. Not being able to make a discrimination, between the nodes **405e** and **405g**, then the node **405f** looks which, among the nodes **405e** and **405g**, has elected the lowest number of MPR nodes: the node **405e** has elected as MPR nodes the nodes **405a** and **405f**, whereas the elected MPR nodes of the node **405g** are the nodes **405f** and **405h**; again, it is not possible to discriminate between the nodes **405e** and **405g**, so the DFC client resident on the node **405f** performs a random choice, and for example the choice is made of node **405g**.

Let now the case be considered that loops exist. In the worst case, for each DFA client **120** there is only one DFC client **215** that can control it. It is therefore necessary to avoid that two DFC clients **215** choose to control a same DFA client **120**, otherwise it may happen that the control of (the DFA client **120** resident on) all the nodes in the subset **450** of elected MPR nodes is not guaranteed. In this case, the DFC client **215** should select the DFA client **120** to control following a certain control direction, denoted

610 in the example of Figure 6. In particular, according to an embodiment of the present invention, the control direction is defined by the node, in the loop, having the numerically lowest IP address; for example, making reference to Figure 6, and assuming that the IP addresses of the nodes 605a, 605b, 605c, 605d and 605e are *210*, *212*, *214*, *216* and *218*, the numerically lowest address is *210*. In the *TC* message 500 transmitted by the node 605a, each DFC client 215 identifies the first node 505a announced in the MPR selector set among those making part of the loop. Thus, each DFC client 215 chooses, as DFA client 120 to control, the one residing on the MPR node that, starting from the node where the considered DFC client 215 resides and passing through the node in the loop having lowest IP address, allows reaching the first node 505 announced in the MPR selector set in the *TC* message. Referring to the example of Figure 6, a loop 600 exists formed by the nodes 605a, 605b, 605c, 610d and 610e; in this case, the node of the loop 600 having the lowest IP address is the node 605a (IP address = *210*), so this node 605a is the one defining the control direction to be followed. All the DFC clients 215 resident on the nodes of the loop 600 look at the MPR selector set announced by the node 605a in the *TC* messages it sends; in this example, the MPR selector set is assumed to be formed by the nodes 605e and 605b, and it is also assumed that the node 605e is the first to be announced in the MPR selector set (this node is also part of the loop). Thus, the control direction 610 to be followed is the direction that, passing through the node 605a (*i.e.* the node with lowest IP address in the loop), allows reaching the node 605e (the first announced node in the MPR selector set). The DFC client 215 resident on the node 605c thus chooses to control the DFA client 120 resident on the node 605b, which allows it to reach the node 605e passing through the node 605a; similarly, the DFC client 215 resident on the node 605d chooses to control the DFA client 120 resident on the node 605c, which allows it reaching the node 605e passing through the node 605a. In this way, each node of the loop is controlled by another node.

Hereinafter, examples of countermeasures that can be taken when it is detected a violation of the security policies will be described.

Local repair mechanism

The local repair mechanism calls for placing in quarantine the network node where a DFA client 120 resides which has not properly enforced the security policies defined for the MANET. The quarantine has effect locally to the MANET (*i.e.* the DFC server 205 is not involved) and it is directed to temporarily isolate the detected node on which the DFA client 120 which has not properly enforced the security policies resides.

In particular, in an embodiment of the present invention, the quarantine mechanism is applied to a generic node a maximum, predetermined number of times, defined for example by a threshold value; once a node has been put in quarantine the maximum number of times, a global repair mechanism (described later on) can be applied to that node.

The choice of allowing a DFA client 120 to be put in quarantine more than once is motivated by

the fact that an incorrect enforcement of the security policies might be caused by a temporary misalignement of a DFA client **120** with the DFA server **110**, due for example to an ongoing update of the security policies, or to a delay of convergence of the routing protocol, thereby the topological indications used for applying the security policies may not be sufficiently up to date.

Every DFC client **215** performs an active control of the DFA client **120** that has been selected for being controlled, by means of a continuous monitoring of the (entering and exiting) traffic flowing through the node hosting the controlled DFA client **120**, so as to detect cases of missed enforcement of the security policies.

Each time a DFC client **215** detects an improper application of the security policies defined for the MANET, it stores data, schematized in **Figure 5B** and denoted therein as **507**, comprising:

- the IP address (IP_S) **510** of the node originating the data packets that violate a certain security policy;
- the IP address (IP_D) **515** of the destination node of those data packets;
- the MAC (Media Access Control) address (MAC_in) **520** of the node originating the incoming traffic entering into the node hosting the controlled DFA client **120**; this MAC address MAC_in corresponds to the node that has routed the traffic to the node hosting the DFA client **120** being controlled;
- the identifier (ID) **525** of the traffic session.

Additionally, every DFC client **215** keeps, for all the MANET nodes, an association between the IP address IP_S of the source node and the related MAC address MAC_S (field **530**).

Based on this data, the DFC client **215** starts the local repair procedure, which calls for:

- identifying the DFA client **120** that has not properly enforced the security policies; and
- put in quarantine the network node on which the identified DFA client **120** resides.

To identify the DFA client **120** that has not properly enforced the security policies, the DFC client **215** exploits the data **507**, particularly the MAC address MAC_in **520** of the traffic entering the controlled DFA client **120**, as well as the routing table; as known to those skilled in the art, the routing table, schematically depicted in **Figure 5C** and denoted therein as **550** is a table that the generic node uses to know how to reach any other network node, and wherein the number of hop counts necessary to reach the other nodes is specified. Using the data contained in the routing table **550**, the DFC client **215** is capable of determining how far is, in terms of hops number ("hop count"), the source node, whose address is the IP_S.

The procedure for identifying the DFA client **120** that has not enforced the security policies is schematized by the flowchart of **Figure 7**.

Firstly, the DFC client searches, in its routing table, the entry corresponding to the node having address equal to IP_S, *i.e.* the source node originating the traffic that violated the security policy, so as to

derive the corresponding hop count value (block **705**).

If the hop count value is equal to 1 (block **710**, exit branch Y), then the node with address IP_S is directly reachable by the node where the DFC client **215** resides. The DFC client **215** checks whether the MAC address MAC_in of the traffic entering the controlled DFA client **120** coincides with the MAC address MAC_S of the source node (block **715**). In the affirmative case (exit branch Y of block **715**), the (node hosting the) controlled DFA client **120** is thus responsible of the security policies violation (*i.e.*, it is the security policy violator node): the DFC client **215**, being the next hop in respect of the security policies violator node, puts the latter node in quarantine (block **720**).

Considering the exemplary situation depicted in **Figure 8A**, let the traffic originated by node **805a** and destined to node **805b** be considered, and let it be assumed that the node **805a** is not authorized, by the security policies defined for the MANET, to generate such traffic. In such a case, the node **805c**, is a direct neighbour of the source node **805a**, and the node **805d** is a direct neighbour of the destination node **805b**. The two nodes **805c** and **805d** are responsible of enforcing the security policies, blocking the traffic generated by the node **805a**. Let it be assumed that the DFA client **120** resident on the node **805c** does not enforce the security policies, and lets the traffic originated by the node **805a** pass through. The DFA client **120** resident on the node **805c** is thus the security policies violator, which should be identified by the DFC mechanism. Let it be assumed that the DFC client **215** resident on a node **805e** is the controller of the DFA client **120** hosted by the node **805c**: inspecting the traffic exiting the node **805c**, the DFC client **215** hosted by the node **805e** is able to detect that the traffic having IP_S equal to that of the node **805a** is not authorized in the MANET. Looking at its routing table, the DFC client **215** hosted by node **805e** determines that the hop count value for the node **805a** is equal to 1, and that the MAC address MAC_in of the traffic entering the controlled DFA client **120** hosted by node **805c** coincides with the MAC address MAC_S of the source node **805a**. Thus, the DFC client **215** hosted by node **805e** determines that the controlled DFA client **120** has not enforced the security policies, and it is a security policies violator: being the controller node **805e** a direct neighbour of the violator node **805c**, the controller node **805e** directly puts the violator node **805c** in quarantine.

Referring back to **Figure 7**, if the two MAC addresses MAC_in and MAC_S do not coincide (exit branch **N** of block **715**), the DFA client **120** that has not enforced the security policies resides on the node that, along the route from the source node (IP address IP_S) to the destination node (IP address IP_D) precedes the controlled DFA client **120**, *i.e.* on the node whose MAC address is MAC_in (in fact, the node responsible of enforcing the security policies is the next hop of the source node, with IP address IP_S, so it cannot be the controlled DFA client **120**, because the latter is distant two hops). The DFC client **215**, not being the next-hop of the security policies violator node, does not directly put the violator node in quarantine: it instead broadcasts an alarm signal, particularly a *NOTIFICATION* message (block **725**), described in detail later on, for notifying to the other MANET nodes that the identified DFA client **120** has

not properly enforced the security policies, for asking the neighbouring nodes to put the violator node in quarantine, and for asking the nodes that are along the route from the source node to the destination node to temporarily apply the violated security policies.

Considering the exemplary situation depicted in **Figure 8B**, let the traffic originated by the source node **805a** and destined to the destination node **805b** be considered, assuming that the node **805a** is not authorized, by the security policies defined for the MANET, to generate such traffic. Node **805c** is a direct neighbour of the source node **805a**, and node **805d** is a direct neighbour of the destination node **805b**; the two nodes **805c** and **805d** are thus responsible of enforcing the security policies, blocking the traffic generated by the node **805a**. Let it be assumed that the DFA client **120** resident on node **805c** does not enforce the security policies, and lets the traffic originated by node **805a** pass through; the node **805c** is thus the security policies violator, that should be identified by the DFC mechanism. The DFC client **215** resident on node **805f** is the controller of the DFA client **120** resident on node **805g**. Inspecting the traffic leaving node **805g**, the DFC client **215** resident on node **805f** detects that the traffic source IP address IP_S corresponding to the IP address of the source node **805a** is not authorized in the MANET. Based on its routing table, the DFC client **215** resident on node **805f** derives that the hop count value for the source node **805a** is equal to 1, but the MAC address MAC_in of the traffic entering the controlled DFA **120** client resident on node **805g** differs from the MAC address MAC_S of the source node **805a**. Thus, the DFA client **120** that has not enforced the security policies resides on the node having MAC address equal to MAC_in, *i.e.* on the node **805c**. The DFC client **215** resident on the node **805f** cannot directly put in quarantine the violator node **805c**, thus it broadcasts a *NOTIFICATION* alarm message.

Referring back again to **Figure 7**, if the hop count is not equal to 1 (exit branch **N** of block **710**), the DFC client **215** ascertains whether the hop count value is equal to 2 (block **730**).

If the hop count value is equal to 2 (exit branch **Y** of block **730**), then the source node originating the traffic in violation of the security policies is not directly reachable by the DFC client **215**. The DFC client **215** then checks if the MAC address MAC_in of the traffic entering the controlled DFA client **120** coincides with the MAC address MAC_S of the source node (block **735**). In the affirmative case (exit branch **Y** of block **735**), the DFA client **120** controlled by the considered DFC client **215** is responsible of the security policies violation, *i.e.* it is the violator node. The DFC client **215**, being in this case a next hop of the violator node, directly puts the violator node in quarantine (block **720**).

Referring to the exemplary scenario depicted in **Figure 8C**, let the traffic originated by source node **805a** and destined to the destination node **805b** be considered, and let it be assumed that the source node **805a** is not authorized, based on the security policies valid for the MANET, to generate such traffic. In this situation, node **805c** is a direct neighbour of the source node **805a**, and node **805d** is a direct neighbour of the destination node **805b**. The two nodes **805c** and **805d** are thus responsible of the enforcement of the security policies, blocking the traffic generated by the source node **805a**. Let it be

assumed that the DFA client **120** resident on the node **805c** does not enforce the security policies, letting the traffic originated by the source node **805a** pass through: the (DFA client **120** on) node **805c** is thus the security policies violator, that should be identified by the DFC mechanism. The DFC client **215** resident on node **805h** is one of the controllers of the DFA client **120** on node **805c**. Inspecting the traffic exiting the node **805c**, the DFC client **215** resident on node **805h** detects that the traffic with source IP address IP_S equal to the IP address of the source node **805a** is not authorized in the MANET. Looking at its routing table, the DFC client **215** resident on node **805h** derives that the hop count value for the source node **805a** is equal to 2, and that the MAC address MAC_in of the traffic entering the controlled node **805c** coincides with the MAC address MAC_S of the source node **805a**. Thus, the DFA client **120** resident on node **805c** controlled by the DFC client **215** resident on node **805h** has not enforced the security policies, and node **805c** is the violator; since the controller node **805h** is a next-hop of the violator node **805c**, the DFC client **215** resident on node **805h** directly puts the violator node **805c** in quarantine.

Back to **Figure 7**, if the two MAC addresses MAC_in and MAC_S do not coincide (exit branch **N** of block **735**), the DFC client **215** checks if the IP address IP_S of the source node is included in the MPR selector set announced in the *TC* messages issued by the node having the MAC address MAC_in of the traffic entering the controlled node (block **740**).

In the affirmative case (exit branch **Y** of block **740**), the security policy violator node is the node which, along the route from the source node (with IP address IP_S) to the destination node (with IP address IP_D), precedes the controlled node, *i.e.* the node having MAC address equal to MAC_in. In such a case, the considered DFC client **215**, not being the next hop of the violator node, cannot directly put the latter in quarantine, so the DFC client **215** broadcasts a *NOTIFICATION* alarm message (block **725**), concerning the violator node.

Referring to the exemplary scenario shown in **Figure 8D**, let it be assumed that there is a traffic flow originating from the source node **805a** and directed to the destination node **805b**, assuming that the node **805a** is not authorized, by the security policies defined for the MANET, to originate such traffic. Node **805c** is a direct neighbour of the source node **805a**, and node **805d** is a direct neighbour of the destination node **805b**; the nodes **805c** and **805d** are thus responsible of enforcing the security policies, blocking the traffic generated by the source node **805a**. It is assumed that the DFA client **120** resident on node **805c** does not enforce the security policies, letting the traffic coming from the source node **805a** pass through; the node **805c** is thus a security policies violator that has to be identified by the DFC mechanism. The DFC client **215** resident on node **805i** is the controller of DFA client **120** resident on the node **805m**. Inspecting the traffic exiting the node **805m**, the DFC client **215** resident on node **805i** detects that the traffic with IP address corresponding to the address IP_S of the source node **805a** is not authorized. Based on its routing table, the DFC client **215** determines that the hop count value for the source node **805a** is equal to 2, and that the MAC address MAC_in of the traffic entering the node **805m** does not

coincide with the MAC address MAC_S of the source node **805a**, being instead the MAC address of the node **805c**. The controller node **805i** checks whether, in the *TC* messages sent by the node **805c**, the IP address IP_S of the source node **805a** is included in the MPR selector set: thus, the DFA client **120** that has not enforced the security policies (the security policies violator) resides on node **805c**. In such a case, the node **805i**, not being a direct neighbour of the security policies violator node, broadcasts a *NOTIFICATION* alarm message.

Back to **Figure 7**, if the IP address IP_S of the source node originating the traffic is not included in the MPR selector set (exit branch **N** of block **740**) announced by the node with MAC address equal to MAC_in, then the DFC client **215** ascertains whether its MAC address coincides with the MAC address MAC_in of the traffic entering the controlled DFA client **120** (block **745**). If the two addresses coincides (exit branch **Y** of block **745**), the DFA client **120** that has not enforced the security policies resides on the node that precedes the node hosting the DFC client **215** along the route from the source node (IP address IP_S) to the destination node (IP address IP_D), and thus said DFA client **120** resides on a neighbouring node (a next hop) of the node where the DFC client **215** resides. In order to identify which of its neighbours has not enforced the security policies, the DFC client **215** searches, in the MPR selector set included in the *TC* messages sent by its neighbours, the IP address IP_S of the source node (block **750**). Then, if the IP address IP_S is found, the DFC client **215** checks how many of its neighbours announce the IP address IP_S in their MPR selector set: if only one of the neighbours announces the IP address IP_S in its MPR selector set (exit branch **Y** of block **755**), then such neighbour is the node where the DFA client **120** that violated the security policies resides: the DFC client **215**, residing on a node that is a next-hop of the violator node, directly puts the latter node in quarantine (block **720**); if instead more than one of the neighbour nodes announces the IP address IP_S, then the DFC client **215**, not being able to identify which is the node where the DFA client **120** that violated the security policies resides, broadcasts an alarm message, particularly an *ALERT* message (block **760**).

Considering the exemplary scenario depicted in **Figure 8E**, let the traffic originated by the source node **805a** and destined to node **805b** be considered, assuming that the node **805a** is not authorized, by the security policies defined for the MANET, to generate such traffic. Node **805c** is a direct neighbour of the source node **805a**, and node **805d** is a direct neighbour of the destination node **805b**. The two nodes **805c** and **805d** are thus responsible of enforcing the security policies, blocking the traffic generated by the node **805a**. Let it be assumed that the DFA client **120** resident on node **805c** does not enforce the security policies, letting the traffic originated by node **805a** pass through: the DFA client **120** resident on node **805c** is thus the security policies violator, that should be identified by the DFC mechanism. The DFC client **215** resident on node **805n** is the controller of the DFA client **120** resident on node **805p**; monitoring the traffic on node **805p**, the DFC client **215** resident on node **805n** detects that the traffic having IP address equal to the IP address IP_S of the source node **805a** is not authorized. Based

on its routing table, the DFC client **215** derives that the hop count for node **805a** is equal to 2, and that the MAC address MAC_in of the traffic entering the controlled DFA client **120** resident on node **805p** coincides with its own MAC address. In this case, the controller node **805n** can decree that the DFA client **120** violating the security policies resides on one of its neighbours, particularly on the preceding node along the route from the source node to the destination node. In order to identify which is the violator node, among its neighbours **805q**, **805c**, **805r**, the DFC client **215** resident on the controller node **805n** checks whether, in the *TC* messages sent by the neighbour nodes, there is the IP address IP_S of the source node **805a**. The nodes **805q** and **805r**, not being MPR nodes, do not send *TC* messages (in the drawing, the nodes that have been elected as MPR nodes are encircled). In the considered example, only the node **805c** sends *TC* messages containing the IP address IP_S of the source node **805a**, thus the DFA client **120** that has not enforced the security policies resides on node **805c**; the node **805n**, being a direct neighbour thereof, directly puts the violator node **805c** in quarantine.

Back to **Figure 7**, if the hop count is different from 2 (exit branch **N** of block **730**), the DFC client **215** checks whether the hop count is equal to 3 (block **765**).

In the affirmative case (exit branch **Y** of block **765**), the DFA client **120** that violated the security policies is not the one controlled by the considered DFC client **215**. The DFC client **215** checks if the IP address IP_S of the source node is included in the MPR selector set announced in the *TC* messages sent by the node having MAC address equal to MAC_in, *i.e.* the MAC address of the traffic entering the DFA client **120** controlled by the DFC client **215** considered (block **770**). In the affirmative case (exit branch **Y** of block **770**), the DFA client **120** that violated the security policies resides on the node that, along the route from the source node (IP address equal to IP_S) to the destination node (IP address equal to IP_D), precedes the node hosting the controlled DFA client **120**, *i.e.* the violator node is the node having MAC address equal to MAC_in. In this case, the DFC client **215** is not a next-hop of the violator node, and cannot directly put the violator node in quarantine; the DFC client **215** thus broadcasts an alarm message, particularly a *NOTIFICATION* message (block **725**) related to the violator node.

Looking at the exemplary scenario depicted in **Figure 8F**, let the traffic originating from source node **805a** and destined to destination node **805b** be considered, and let it be assumed that the source node **805a** is not authorized, by the security policies defined for the MANET, to generate such traffic. Node **805c** is a direct neighbour of the source node **805a**, and node **805d** is a direct neighbour of the destination node **805b**; the two nodes **805c** and **805d** are responsible of enforcing the security policies, blocking the traffic generated by the node **805a**. It is assumed that the DFA client **120** resident on node **805c** does not enforce the security policies, letting the traffic generated by the source node **805a** pass through: the node **805c** is thus the security policies violator, that should be identified by the DFC mechanism. The DFC client **215** resident on node **805r** is the controller of the DFA client **120** resident on node **805s**. Monitoring the traffic on node **805s**, the DFC client **215** resident on node **805r** detects that the traffic with IP address

equal to that (IP_S) of the source node **805a** is not authorized; based on its routing table, the DFC client **215** resident on node **805r** determines that the hop count for the source node **805a** is equal to 3; the DFC client **215** also checks if the IP address of the source node **805a** is included in the MPR selector set announced in the *TC* messages sent by the node having MAC address MAC_in, from which the traffic entering the controlled node **805s** comes. The violator node **805c**, having MAC address MAC_in, is thus the node that precedes the controlled node **805s**. The DFC client **215** resident on node **805r** is not a next-hop of the violator node **805c**, thus it cannot directly put the violator node in quarantine: the node **805r** then broadcasts an alarm message, particularly a *NOTIFICATION* message related to the violator node **805c**.

Back to **Figure 7**, if the IP address IP_S of the source node is not included in the MPR selector set announced in the *TC* messages of the node having MAC address equal to the MAC address MAC_in of the traffic entering the DFA client **120** controlled by the DFC client **215** considered (exit branch **N** of block **770**), the DFC client **120** is not able to identify the violator node, and thus simply issues an *ALERT* message (block **760**).

If the hop count is higher than 3 (exit branch **N** of block **765**), the DFA client **120** that did not enforce the security policies is not that resident on the node controlled by the DFC client **215**, neither does it reside on the node that precedes the node where the DFA client **120** controlled by the considered DFC client **215** resides, along the route from the source node to the destination node. In this case, the DFC client **215**, not being able to determine which is the DFA client **120** that violated the security policies, performs no control on the MAC address MAC_in, and directly broadcasts an alarm message, particularly an *ALERT* message (block **760**). However, the DFA client **120** that did not enforce the security policies has in this case already been identified by the monitoring action performed by the nodes distant 1 or 2 hops from the source node; in fact, the filtering is performed only by the neighbour nodes of the node originating the unauthorized traffic. The control of the nodes more distant than 2 hops, which are not responsible of enforcing the security policies in respect of the source node, is scarcely useful for identifying the DFA client **120** that did not enforce the security policies. Such a control is nevertheless useful to detect unauthorized traffic in the MANET, and thus to alert the nodes along the route from the source node to the destination node, avoiding that such unauthorized traffic is routed further until these nodes receive a *NOTIFICATION* message (or a *QUARANTINE* message, as described later) generated as a consequence of the control performed by the DFC clients **215** controlling the DFA clients **120** resident on nodes that are closer to the node originating the unauthorized traffic.

Quarantine

As explained in the foregoing, each time a DFC client **215** detects or is informed by other DFC clients **215** that the DFA client **120** resident on a direct neighbouring node violates the security policies, the DFC client **215** puts the policy violator node in quarantine.

The schematic flowchart of **Figure 9** depicts the quarantine procedure, in an embodiment of the present invention.

Firstly, the DFC client **215** increments a first counter (hereinafter referred to as "FW counter") that is used to keep track of the number of consecutive times that the considered DFA client **120**, to be put in quarantine, has not properly enforced the security policies defined for the MANET (block **905**).

The DFC client **215** then checks if the value of the FW counter for the considered DFA client **120** has reached a predetermined threshold FwThr (block **910**); in the affirmative case (exit branch Y of block **910**), the DFC client **215** starts the global repair procedure in respect of that DFA client **120** (block **915**), as described in detail later on; in such a case, the DFC client **215** preferably broadcasts a *QUARANTINE* message with a flag set to indicate that the global repair procedure has been started for that DFA client **120**, so as to avoid that different DFC clients **215** start the global repair procedure for the same DFA client **120**.

Then, the DFC client **215** increments a second counter (hereinafter referred to as the "source counter"), that is used to keep track of the number of consecutive times that the source node (having IP address IP_S) originating the traffic that violates the security policies has attempted to generate unauthorized traffic (block **920**). The DFC client **215** checks (block **925**) if the value of the source counter has reached a predetermined threshold value SThr: in the affirmative case (exit branch Y of block **925**), the DFC client **215** starts the global repair procedure (block **930**) for the source node having IP address IP_S, and sets the respective flag to be included in a *QUARANTINE* message that is successively broadcast, so as to avoid that other DFC clients **215** start the global repair procedure in respect of that node.

The quarantine involves temporarily isolating the node hosting the DFA client **120** that did not enforce the security policies. The node hosting the DFC client **215** ceases announcing the violator node as its neighbour in *HELLO* messages it sends (block **935**), and activates a third counter, hereinafter referred to as "Qtimer" (block **940**). Then, the DFC client **215** broadcasts a *QUARANTINE* message (block **945**), so that the other network nodes are made aware of the security policies violation, and remove from their routing tables the routes involving the violator node. Until the value of the timer Qtimer reaches the prescribed threshold value (exit branch N of block **950**), the DFC client **215** continues to broadcast *HELLO* messages without announcing the violator node (**955**). When instead the value of the timer Qtimer reaches the preset threshold value (exit branch Y of block **950**), the DFC client **215** re-inserts the node hosting the violator DFA client **120** into the *HELLO* messages it broadcasts (block **960**), and resets the time Qtimer (block **965**): the quarantine is thus stopped, and the DFA client **120** that violated the security policies is re-inserted in the routing tables of the MANET nodes.

In the schematic flowchart of **Figures 10A** and **10B** the management of the alarm messages that have been mentioned in the foregoing and that are used in an embodiment of the present invention is

depicted.

As mentioned in the foregoing, the *NOTIFICATION* message is broadcasted to all the MANET nodes whenever a DFC client **215**, during the monitoring of the traffic on the controlled DFA client **120**, detects a violation of the security policies defined for the MANET, and identifies the violator DFA client **120** that has not enforced the security policies, but the DFC client **215** cannot directly put the violator node in quarantine, because the violator node is not a next-hop thereof.

The purpose of the *NOTIFICATION* message is:

- notifying to the other MANET nodes that a certain DFA client **120** has not properly enforced the security policies;
- ask the neighbouring nodes to put the violator node in quarantine; and
- ask the nodes that are along the route from the source node to the destination node to temporarily apply the violated security policies.

The structure of the *NOTIFICATION* message is depicted in **Figure 11**; the message includes:

- a field **1105** for carrying the IP address IP_S of the source node originating the data packets violating the security policies;
- a field **1110** for carrying the IP address IP_D of the destination node of the data traffic;
- a field **1115** for carrying the IP address IP_DFA of the node where the DFA client **120** that has not enforced the security policies resides;
- a field **1120** for carrying an identifier ID of the traffic session;
- a field **1125** for carrying an identifier RuleID of the security policy (firewalling rule) that has been violated; and
- a field **1130** for carrying the IP address IPNotification of the node where the DFA client **120** resides that is controlled by the DFC client **215** and that has detected traffic not compliant with the security policy identified by the identifier RuleID.

Referring to **Figure 10A**, each DFC client **215**, upon receiving an alarm message, checks (block **1005**) if the received message is a message of the type *NOTIFICATION*. In the affirmative case (exit branch **Y** of block **1005**), the DFC client **215** checks if the node where it resides is a direct neighbour of the node where the violator DFA client **120** resides (the node whose IP address IP_DFA is contained in the field **1115**) (block **1010**): in the affirmative case (exit branch **Y** of block **1010**), the DFC client **215** checks whether it has already put the violator node in quarantine (block **1015**). If the violator node has already been put in quarantine (exit branch **Y** of block **1015**), the DFC client **215** stops propagating the *NOTIFICATION* message (block **1020**). If on the contrary the violator node has not been put in quarantine yet (exit branch **N** of block **1015**), the DFC client **215** starts the quarantine procedure in respect of the violator node (block **1025**) and then it stops propagating the *NOTIFICATION* message (block **1020**).

If the node where the DFC client **215** that received the *NOTIFICATION* message is not a direct

neighbour of the violator node (exit branch **N** of block **1010**), then the DFC client **215** checks whether the node where it resides is on the route from the source node (IP address IP_S) to the destination node (IP address IP_D), by assessing if the IP address of the node where it resides coincides with the address IPNotification contained in the field **1130** of the received message (block **1030**). In the affirmative case (exit branch **Y** of block **1030**), the DFC client **215** assesses whether it has already applied the security policy (firewalling rule) that has been violated (identified by the identifier RuleID contained in the field **1125** of the received message) (block **1035**). If it has (exit branch **Y** of block **1035**), then it means that the DFC client **215** has already received previous *NOTIFICATION* alarm messages, possibly by other DFC clients **215** that control it, and thus it stops further propagating the *NOTIFICATION* message (block **1020**). Differently (exit branch **N** of block **1035**), it applies the violated security policy (identified by the identifier Rule ID) (block **1040**), and stops routing further data packets of the type identified by the identifier ID contained in the field **1120** of the received message. The DFC client **215** then checks (block **1045**) if it already received previous *NOTIFICATION* messages in respect of the same traffic session identifier by the identifier ID: in the negative case (exit branch **N** of block **1045**), it propagates further the message (block **1050**), otherwise (exit branch **Y** of block **1045**) it stops propagating the message (block **1020**).

If the DFC client **215** assesses that the node where it resides is not on the route from the source node to the destination node (exit branch **N** of block **1030**), then it checks whether it already received previous *NOTIFICATION* messages for the same traffic session (block **1045**); in the affirmative case (exit branch **Y** of block **1045**), it stops propagating the message (block **1020**), otherwise (exit branch **N** of block **1045**), it further propagates the message (block **1050**).

The *QUARANTINE* message is broadcasted to all the MANET each time a DFC client **215** detects that a controlled DFA client **120** has not enforced the security policies, and, being the node where it resides a next-hop of the node where the violator DFA client **120** resides, it directly puts the latter in quarantine. The *QUARANTINE* message is also broadcasted by a DFC client **215** that, having received a *NOTIFICATION* message, recognizes to be a direct neighbour of the violator DFA client **120**, and thus puts the latter in quarantine.

The purpose of the *QUARANTINE* message is that of informing the other nodes that a certain node has been put in quarantine, and thus it allows keeping the counters FW counter and source counter aligned.

The structure of the *QUARANTINE* message is depicted in **Figure 12**; the message includes:

- a field **1205** for carrying the IP address IP_S of the source node originating the data packets violating the security policies;

- a field **1210** for carrying the IP address IP_D of the destination node of the data traffic;

- a field **1215** for carrying the IP address IP_DFA of the node where the DFA client **120** that has not enforced the security policies resides;

- a field **1220** for carrying an identifier ID of the traffic session;

- a field **1225** for carrying an identifier RuleID of the security policy (firewalling rule) that has been violated;

- a field **1230** for carrying a flag CGR ("Client Global Repair") used to notify to the other nodes that a global repair procedure has been started for the violator DFA client **120** residing on the node of address IP_DFA; and

- a field **1235** for carrying a flag SGR ("Source Global Repair") used to notify to the other nodes that the global repair procedure has been started for the source node of address IP_S that generated the unauthorized traffic.

Referring to **Figure 10B**, the DFC client **215**, upon verifying that received alarm message is not a *NOTIFICATION* message (exit branch **N** of block **1005**), checks whether the message is a *QUARANTINE* message (block **1055**). In the affirmative case (exit branch **Y** of block **1055**) it checks whether the node where it resides is a direct neighbour of the node where the violator DFA client **120** resides, identified by the address IP_DFA contained in the field **1215** of the received message (block **1060**): if this is not the case (exit branch **N** of block **1060**), the DFC client **215** checks whether it already received previous *QUARANTINE* messages for the same session identified by the session identifier ID contained in the field **1220** of the received message (block **1065**), in which case (exit branch **Y** of block **1065**) it stops propagating the *QUARANTINE* message (block **1070**); if no *QUARANTINE* messages were previously received (exit branch **N** of block **1065**), the DFC client **215** increases the counters FW counter and source counter (block **1075**), stops forwarding the *NOTIFICATION* messages (block **1080**) received/generated in respect of the same session identified by the identifier ID, and propagates the *QUARANTINE* message to the other nodes (block **1081**).

If the DFC client **215** instead realizes to be a direct neighbour of the node where the violator DFA client **120** resides (exit branch **Y** of block **1060**), it checks whether the violator node has already been put in quarantine (block **1083**): if it has not (exit branch **N** of block **1083**), the DFC client **215** puts the violator node in quarantine (connector **B** and block **1025**), otherwise (exit branch **Y** of block **1083**) it stops propagating the *QUARANTINE* message (block **1070**).

ALERT alarm messages, differently from the *NOTIFICATION* and *QUARANTINE* messages, are sent only to the next hops of the node hosting a DFC client **215** each time the latter is more than 3 hops far from the source node of the unauthorized traffic, and only in case the DFC client **215** is not able to identify which is the DFA client **120** that has not enforced the security policies. The *ALERT* message is sent to inform a DFA client **120** controlled by the DFC client **215** about the presence of unauthorized traffic along the route from the source node to the destination node, so that the DFA client **120**can block the unauthorized traffic.

The structure of the *ALERT* message is depicted in **Figure 13**; the message includes:

- a field **1305** for carrying the IP address IP_S of the source node originating the data packets violating the security policies;

- a field **1310** for carrying the IP address IP_D of the destination node of the data traffic;

- a field **1315** for carrying the IP address IP_DFA of the node where the DFA client **120** that has not enforced the security policies resides;

- a field **1320** for carrying an identifier ID of the traffic session;

- a field **1325** for carrying an identifier RuleID of the security policy (firewalling rule) that has been violated;

- a field **1330** for carrying the IP address IP_Alert of the node hosting the DFA client **120** controlled by the DFC client **215** and that detected the presence of unauthorized traffic.

Upon receiving an *ALERT* message (exit branch **N** of block **1055**), the DFC client **215** checks whether it is on the route from the source node to the destination node, by comparing its own IP address with the address IP_Alert contained in the field **1330** of the received message (block **1085**). If the two addresses coincide (exit branch **Y** of block **1085**), then the DFC client **215** ascertains whether it has already applied the violated security policies (identified by the identifier RuleID) (block **1087**). If not (exit branch **N** of block **1087**), the security policy is applied (block **1089**), otherwise it means that previous *ALERT* messages were already received for the same session, and the message is thus ignored (block **1091**). In both cases, the DFC client **215** waits for the next *NOTIFICATION* and *QUARANTINE* messages (block **1093**). A DFC client **215** that, upon receiving the *ALERT* message, assesses not to be along the route from source node to the destination node (exit branch **N** of block **1085**) ignores the *ALERT* message (block **1091**) and keeps on listening next *NOTIFICATION* and *QUARANTINE* messages (block **1093**).

Global repair mechanism

The global repair mechanism involves the intervention of the DFC server **205**, which is responsible of centrally managing the DFC mechanism, each time a stronger countermeasure against the security policies violator nodes is requested, in respect of nodes that have repeatedly violated the security policies, not enforcing them properly. The global repair mechanism is used whenever the quarantine did not succeed.

The global repair mechanism may involve the exclusion from the MANET of a node where the violator DFA client **120** resides, identified during the local repair procedure; the node excluded from the MANET may be requested to re-authenticate itself in order to be re-admitted to the MANET.

The global repair procedure in respect of a generic network node is started by the first DFC client **215** whose counter FW counter for that node reaches the predetermined threshold: in such a case, the DFC client **215**, during the local repair procedure, sets the CGR flag in the *QUARANTINE* message it broadcasts. The DFC clients **215** that receive the *QUARANTINE* message with the CGR flag set are informed of the fact that some node already started the global repair procedure, and thus they reset the

respective counters FW counter in respect of the same DFA client **120**.

The global repair procedure involves the transmission to the DFC server **205** of a *GLOBAL REPAIR REQUEST* message, whose structure is pictorially shown in **Figure 14A**. The message, globally denoted **1400**, includes:

- a field **1405** for carrying the IP address IP_S of the source node of the data packets that violate a certain security policy;

- a field **1410** for carrying the IP address IP_DFA of the node where the violator DFA client **120** resides, that has not enforced the security policy; and

- a field **1415** for carrying the identifier RuleID of the violated security policy (firewalling rule).

Upon receiving a *GLOBAL REPAIR REQUEST* message, the DFC server **205** verifies whether a violation actually occurred. This is done by an interaction with the DFA server **110**. The latter:

- checks the correspondence between the security policy identified by the identifier RuleID contained in the received *GLOBAL REPAIR REQUEST* message, and one of the security policies in a list of security policies configured for the MANET;

- ensures that the security policy identified by the identifier RuleID has actually been sent to the source node of the unauthorized traffic and the node where the violator DFA client **120** resides, specified in the received *GLOBAL REPAIR REQUEST* message (the DFA server **110** keeps a list of the active MANET nodes, and, for each of them, an indication of the current security policies (general and specific) applied by the node).

If the DFA server **110** finds that the DFA client **120** identified by the IP address IP_DFA in the received *GLOBAL REPAIR REQUEST* message actually violated the security policies, the DFA server **110** removes the IP address IP_DFA of that node (and thus the firewalling rule defined for it) from its database. Then, the DFA server **110** sends to the authentication server of the MANET the identifier IP_DFA of the node to be de-authenticated. The DFA server **110** also sends to the DFC server **205** a *GLOBAL REPAIR ACCEPT* message, by means of which it requests the exclusion of the node identified by the IP address IP_DFA from the MANET. The structure of the *GLOBAL REPAIR ACCEPT* message is schematically depicted in **Figure 14B**: the message, globally denoted **1450**, contains a field **1455** for carrying the IP address IP_DFA of the node to be excluded from the list of DFC clients **215**, and a field **1460** for carrying a signature of the DFC server **205**. The signature is included in the *GLOBAL REPAIR ACCEPT* message by the DFC server **205**, which then sends it to the node that sent the *GLOBAL REPAIR* REQUEST message; the *GLOBAL REPAIR ACCEPT* message is then broadcasted in the MANET, so that the nodes can exclude the node identified by the IP address IP_DFA, blocking any kind of traffic it generated, exception made for authentication traffic for the re-authentication of the node in the MANET. At the same time, the DFA server **205** informs every node of the exclusion of the DFA client **120** resident on the excluded node; this is done through the periodic polling mechanism by which the DFA clients **120** get from

the DFA server **110** the security policies to be enforced.

The same global repair procedure is also applied when the counter source counter reaches the predetermined threshold.

§ § § § §

In addition to the detection of the improper enforcement of the security policies, the present invention can also be used for detecting and contrasting selfish nodes, which tend to participate to the MANET in such a way that they only route their own data packets, discarding the data packets originating from/directed to other nodes. Exploiting the same mechanisms described in the foregoing (selection of the controlled DFA clients **120** and of the controller DFC clients **215**, local and global repair mechanisms), it is possible to isolate the selfish nodes, thereby promoting cooperation between the nodes.

Referring to **Figure 2**, each node acts as a controller (DFC client **215**) of a certain DFA client **120**, through the constant monitoring of the traffic entering and exiting the controlled DFA client **120**. In order to detect and manage a selfish behaviour of a controlled DFA client **120**, the controller DFC client **215** checks that the packets related to the routing protocol and the application-layer packets are properly routed (complying the security policies defined for the MANET).

In particular, referring to **Figure 15**, for each data packet **1500** entering and/or exiting a controlled DFA client **120**, it is first assessed whether the packet complies with the security policies defined for the MANET (block **1505**); in case of violation of the security policies (exit branch **Y** of block **1505**), this is managed as described in the foregoing, first through the local repair mechanism (block **1510**), and then through the global repair mechanism (block **1520**), if the counters FW counter and/or source counter have reached the predetermined threshold (block **1515**). If the packet complies with the security policies defined for the MANET (exit branch **N** of block **1505**), the selfish node detection logic is applied (block **1525**): for each packet entering the controlled DFA client **120**, having destination IP address IP_D different from the IP address IP_DFA of the node where the controlled DFA client **120** resides, the presence of a corresponding packet exiting the node is checked by the controller DFC client **215**. In case the entering packets are discarded (not let pass through), the controller DFC client **215** decrees a selfish behaviour (exit branch **Y** of block **1525**), and applies a simplified local repair mechanism (**1530**): the controller DFC client **215** increases a dropped packets counter associated with the controlled DFA client **120**; if the dropped packets counter reaches a selfish node detection threshold (set by the controller DFC client **215** according to the security policies of the MANET), then the controller DFC client **215** broadcasts a *SELFISH NODE QUARANTINE* message, schematized in **Figure 16** and including:

- a field **1605** carrying an IP address IP_SELF of the selfish node;
- a field **1610** carrying the MAC address of the selfish node;
- a field **1615** carrying the identifier RuleID of the violated firewalling rule; and
- a field **1620** carrying a flag indicating whether the global repair procedure has already been

started.

Each DFC client **215** which, upon receiving the *SELFISH NODE QUARANTINE* message, realizes to be a direct neighbour of the selfish node, puts the latter in quarantine, stopping to announce the selfish node as its neighbour in the *HELLO* messages, and ignores routing messages coming from the selfish node. In this way, the distant DFC clients, not receiving information about the selfish node, remove it from their routing tables. Similarly to the case of violation of a security policies described above, the quarantine is temporary, and lasts a predetermined time (for example, a timer is used for this purpose), after which the nodes that are direct neighbours of the selfish node restart announcing the selfish node as their neighbour; the selfish node is thus re-inserted in the routing tables of all the nodes of the MANET.

The *SELFISH NODE QUARANTINE* message is also exploited to keep the counters FW counter of the controller DFC clients aligned; these counters are used, as described above, for handling the global repair procedure (a same counter is used for starting the global repair procedure, irrespective of whether a violation of a security policy or a selfish behaviour are detected).

The invention has been here described making reference to an exemplary embodiment thereof, those skilled in the art will recognize that several modifications to the described embodiment, as well as other embodiments are conceivable, without departing from the scope of the invention, defined in the appended claims.

CLAIMS

1. A method of enforcing security policies in a mobile *ad-hoc* network, comprising:

- entrusting at least one first network node, along a data traffic route from a data traffic origin node to a data traffic destination node, with the enforcing of predefined security policies on said data traffic;

- entrusting at least one second network node, distinct from said first network node, with the control of the enforcement of the security policies by said first network node.

2. The method of claim 1, wherein said at least one first network node comprises a network node being a direct neighbour of the origin node.

3. The method of claim 2, wherein said at least one first network node further comprises a network node being a direct neighbour of the destination node.

4. The method of any one of the preceding claims, further comprising:

- providing at least one central security policies distribution server configured to distribute the predefined security policies to the network nodes.

5. The method of any of claims 1 to 4, comprising:

- having the at least one second network node control a compliance with the predefined security policies of the data traffic passing through at least one third network node along said data traffic route.

6. The method of claim 5, further comprising:

having the at least one second network node selecting the said at least one third network node among a set of multipoint relay nodes thereof.

7. The method of claim 6, wherein said having the at least one second network node selecting the at least one third network node comprises selecting the at least one third network node among two or more respective multipoint relay nodes as either:

- the multipoint relay node through which the lowest number of nodes distant two hops can be reached; or

- in case two or more of the multipoint relay nodes of the second network nodes allow reaching a same, lowest number of nodes distant two hops, the multipoint relay node that has in turn elected the lowest number of multipoint relay nodes.

8. The method of claim 7, wherein said having the at least one second network node selecting the at least one third network node further comprises

- in case two or more of the multipoint relay nodes of the second network nodes have elected a same, lowest number of multipoint relay nodes, choosing the at least one third network node at random among said two or more of the multipoint relay nodes of the second network node that have elected the lowest number of multipoint relay nodes.

9. The method of claim 6, 7 or 8, wherein in case the at least one second network node is included in a loop, said selecting the at least one third network node comprises selecting, among the multipoint relay nodes of the second network node, the network node that, starting from the second network node and passing through a node in the loop having lowest network address, allows reaching a first node announced in a multipoint relay selector set in a transmission control protocol message.

10. The method of any one of claims 5 to 9, further comprising:

- in case the second network node detects a security policies violation:

- having the second network node ascertain whether the third network node is the first network node;

- if the third network node is the first network node, having the second network node undertake an action directed to at least temporarily exclude the first network node from the network;

- if the third network node is not the first network node, having the second network node:

- identifying the first network node, and

- in case said first network node is a direct neighbour of the second network node, undertaking an action directed to at least temporarily exclude the first network node from the network;

- in case said first network node is not a direct neighbour of the second network node, or the first network node cannot be identified, broadcasting an alarm message adapted to inform the other network nodes of the security policies violation.

11. The method of claim 10, wherein said at least temporarily exclude from the network comprises causing the first network node to be removed from routing tables of the network nodes in the data traffic routes involving the first network node.

12. The method according to claim 10, further comprising:

- causing the first network node to be de-authenticated from the network after a predetermined number of said temporarily exclude.

13. The method according to any of claims 5 to 12, further comprising:

- having the second network node controlling whether the third network node properly routes data traffic entering thereinto, and

- in the negative case, having the second network node undertake an action to temporarily exclude the third network node from the network.

14. The method of claim 13, wherein said controlling whether the third network node properly routes data traffic entering thereinto comprises:

- having the second network node ascertain that, for each data packet entering the third network node and destined to a destination network node different from the third node, there is a corresponding data packet exiting the third node.

15. A system for enforcing security policies in a mobile *ad-hoc* network, comprising:

- at least one first network node, along a data traffic route from a data traffic origin node to a data traffic destination node, entrusted with the enforcing of predefined security policies on said data traffic;

- at least one second network node, distinct from said first network node, entrusted with the control of the enforcement of the security policies by said first network node.

16. The system of claim 15, wherein said at least one first network node comprises a network node being a direct neighbour of the origin node.

17. The system of claim 16, wherein said at least one first network node further comprises a network node being a direct neighbour of the destination node.

18. The system of any one of claims 15 to 17, further comprising at least one central security policies distribution server configured to distribute the predefined security policies to the network nodes.

19. The system of any of claims 15 to 18, wherein the at least one second network node is configured to control a compliance with the predefined security policies of the data traffic passing through at least one third network node along said data traffic route.

20. The system of claim 19, wherein:

the at least one second network node is adapted to select the at least one third network node among a set of a multipoint relay nodes thereof.

21. The system of claim 20, wherein said at least one second network node is adapted to select the at least one third network node among two or more respective multipoint relay nodes as either:

- the multipoint relay node through which the lowest number of nodes distant two hops can be reached; or

- in case two or more of the multipoint relay nodes of the second network nodes allow reaching a same, lowest number of nodes distant two hops, the multipoint relay node that has in turn elected the lowest number of multipoint relay nodes.

22. The system of claim 21, wherein said at least one second network node is adapted to select the at least one third network node randomly among two or more of the multipoint relay nodes of the second network node that have elected the lowest number of multipoint relay nodes.

23. The system of claim 20, 21 or 22, wherein in case the at least one second network node is included in a loop, the second network node is adapted to select, among the multipoint relay nodes of the second network node, the network node that, starting from the second network node and passing through a node in the loop having lowest network address, allows reaching a first announced node announced in a multipoint relay selector set in a transmission control protocol message.

24. The system of any one of claims 19 to 23, wherein:

- in case the second network node detects a security policies violation, the second network node is adapted to assess whether the third network node is the first network node, and:

- in the affirmative case, to undertake an action directed to at least temporarily exclude the first network node from the network;

- in the negative case:

- to identify the first network node;

- in case said first network node is a direct neighbour of the second network node, undertake an action directed to at least temporarily exclude the first network node from the network;

- in case said first network node is not a direct neighbour of the second network node, or the first network node cannot be identified, to broadcast an alarm message adapted to inform the other network nodes of the security policies violation.

25. The system of claim 24, wherein said at least temporarily exclude from the network comprises causing the first network node to be removed from routing tables of the network nodes in the data traffic routes involving the first network node.

26. The system of claim 25, wherein said at least temporarily excluding from the network further comprises causing the network node responsible of the security policies violation to be de-authenticated from the network after the first network node has been removed a predetermined number times from routing tables of the network nodes in the data traffic routes involving the first network node.

27. The system according to any of claims 19 to 26, wherein the second network node is further adapted to control whether the third network node properly routes data traffic entering thereinto, and, in the negative case to undertake an action to temporarily exclude the third network node from the network.

28. The system of claim 27, wherein the second network node is adapted to ascertain that, for each data packet entering the third network node and destined to a destination network node different from the third node, there is a corresponding data packet exiting the third node.

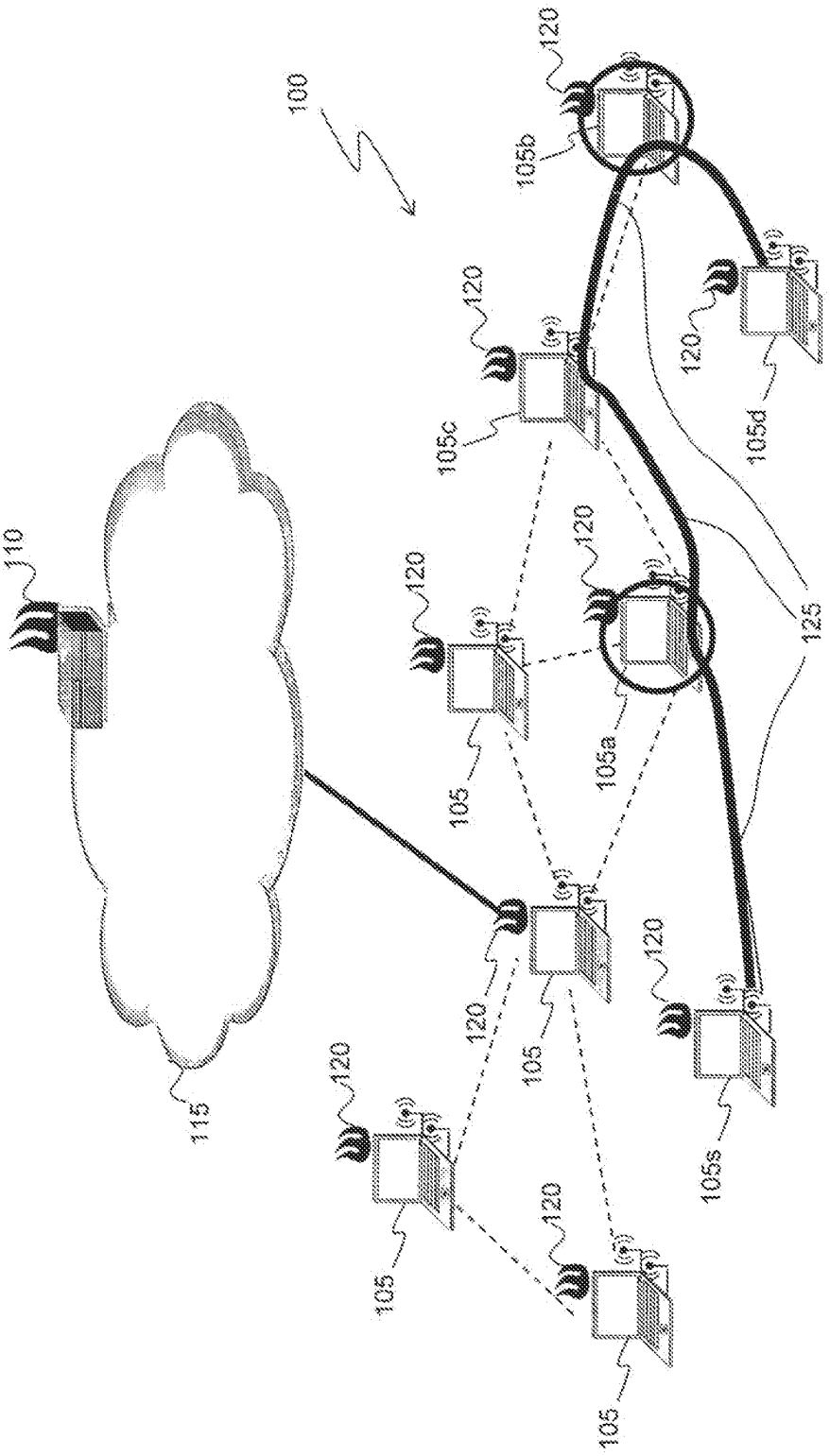


FIG. 1

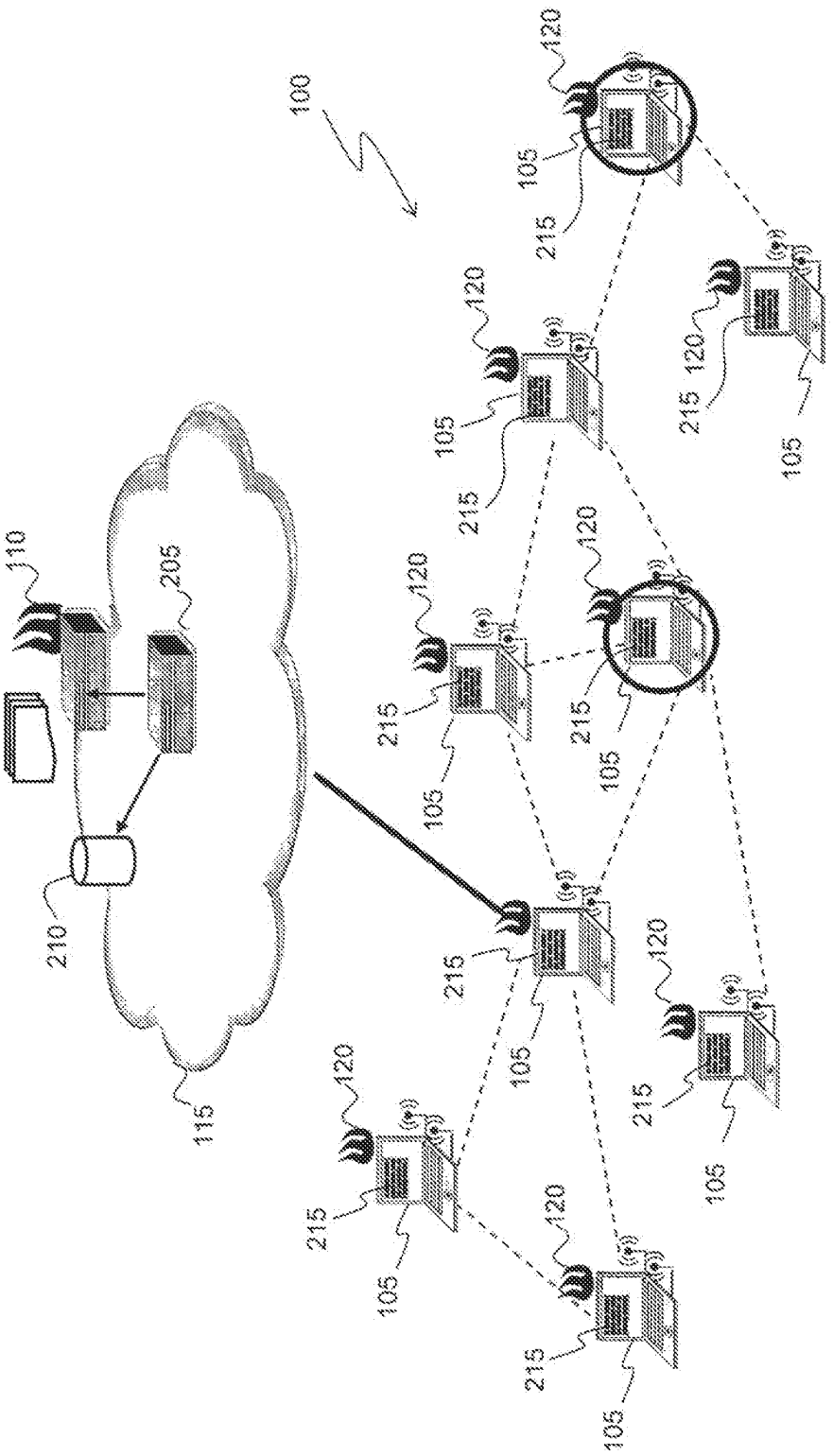


FIG. 2

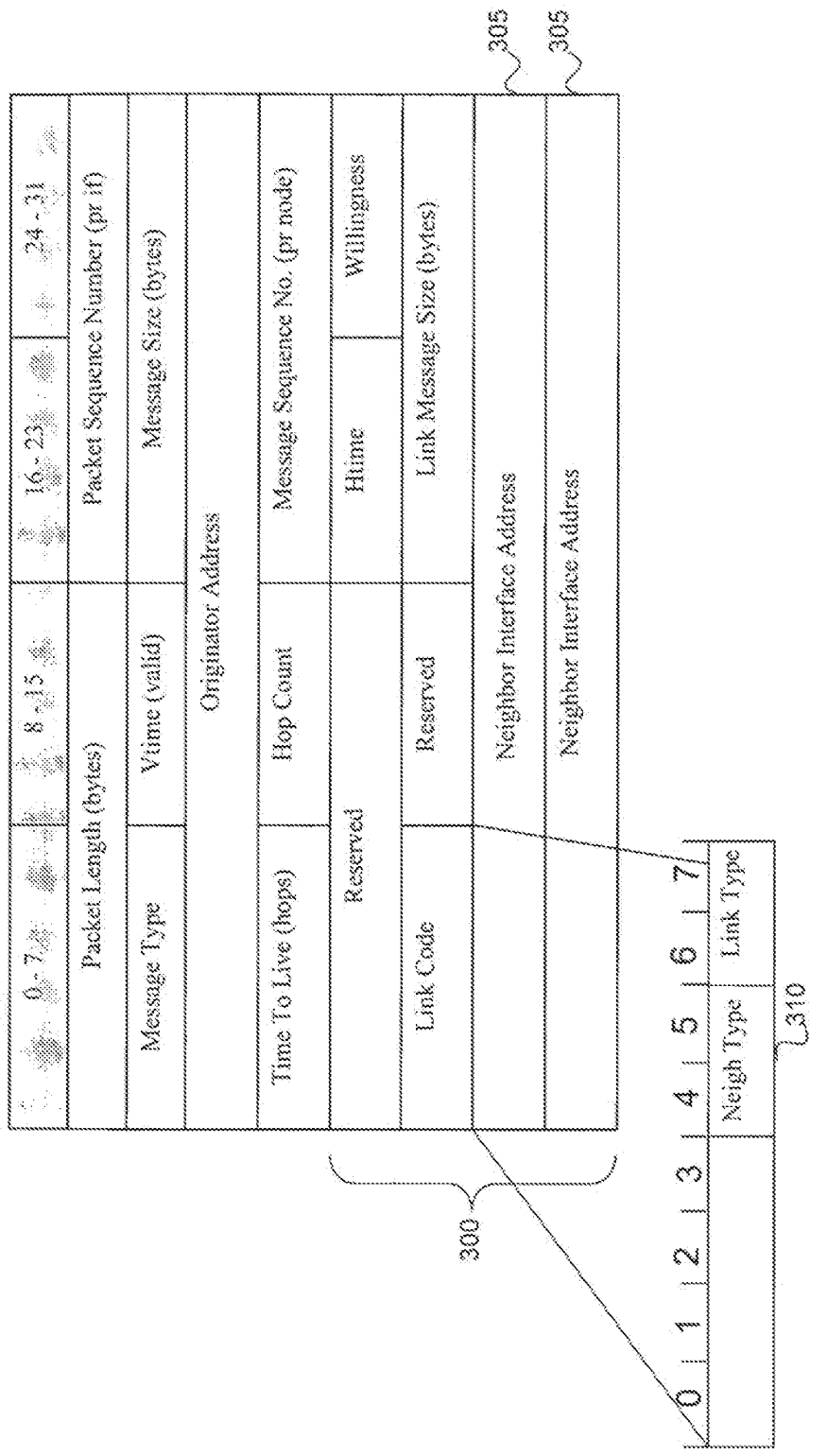

0 - 7
8 - 15
16 - 23
24 - 31
Packet Length (bytes)
Packet Sequence Number (pr if)
Message Type
Vtime (valid)
Message Size (bytes)
Originator Address
Time To Live (hops)
Hop Count
Message Sequence No. (pr node)
Reserved
Htime
Willingness
Link Code
Reserved
Link Message Size (bytes)
Neighbor Interface Address
Neighbor Interface Address
300
305
305
0 1 2 3 4 5 6 7
Neigh Type
Link Type
310

FIG. 3

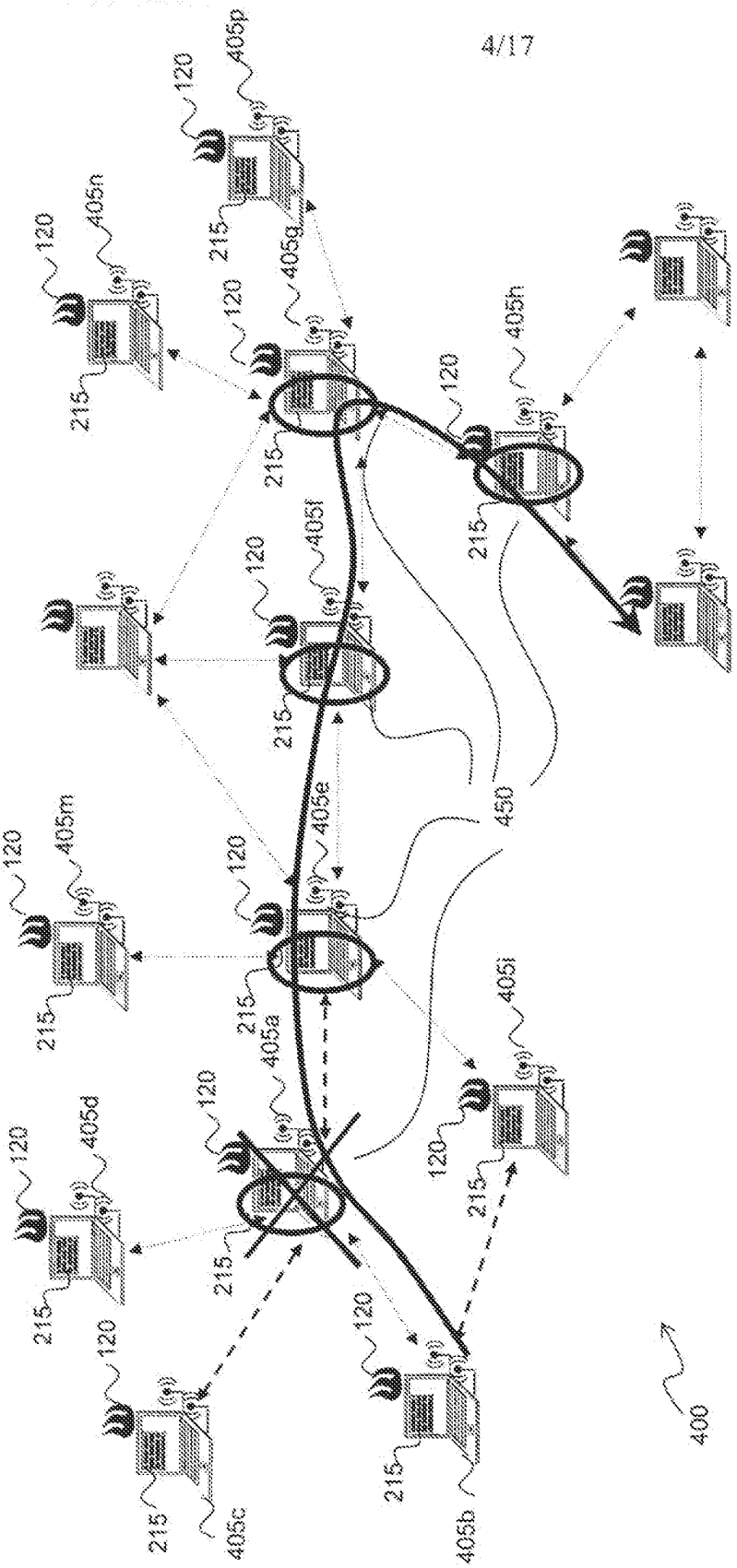

405p
405n
405g
405h
405f
405m
405e
405d
405a
405i
405c
405b
120
215
450
400

FIG. 4

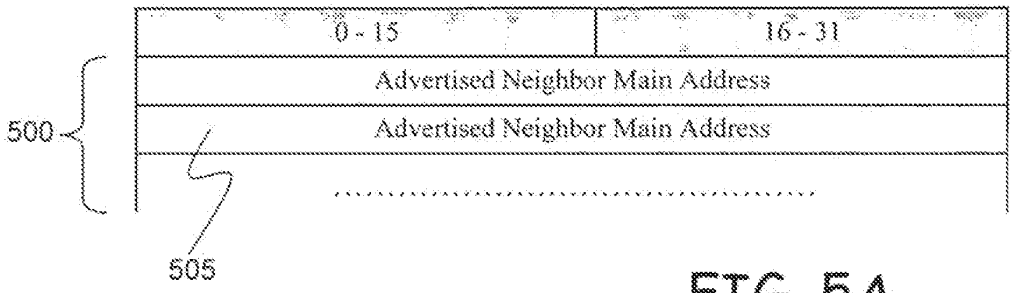


FIG. 5A

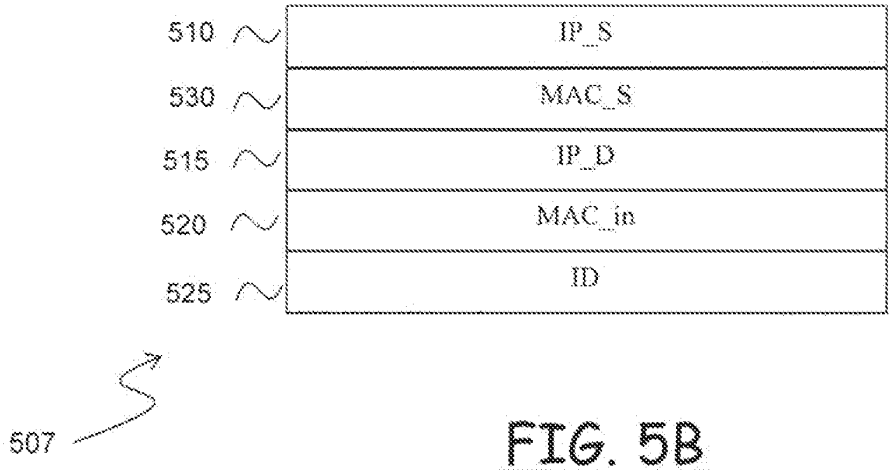


FIG. 5B

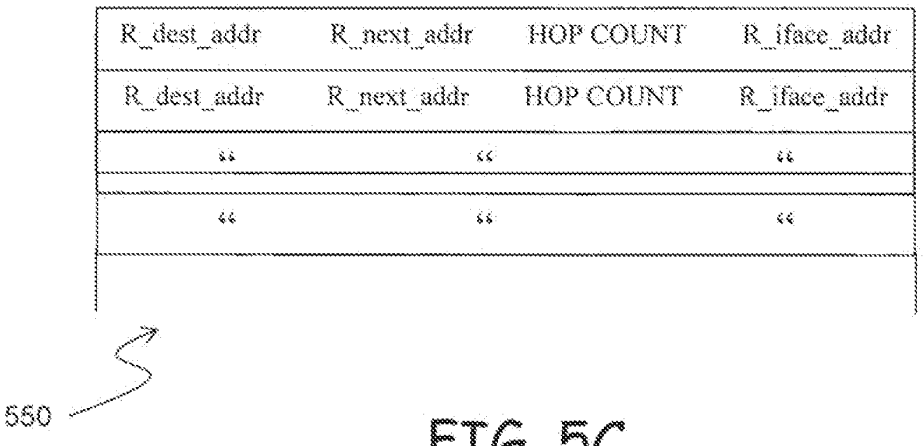


FIG. 5C

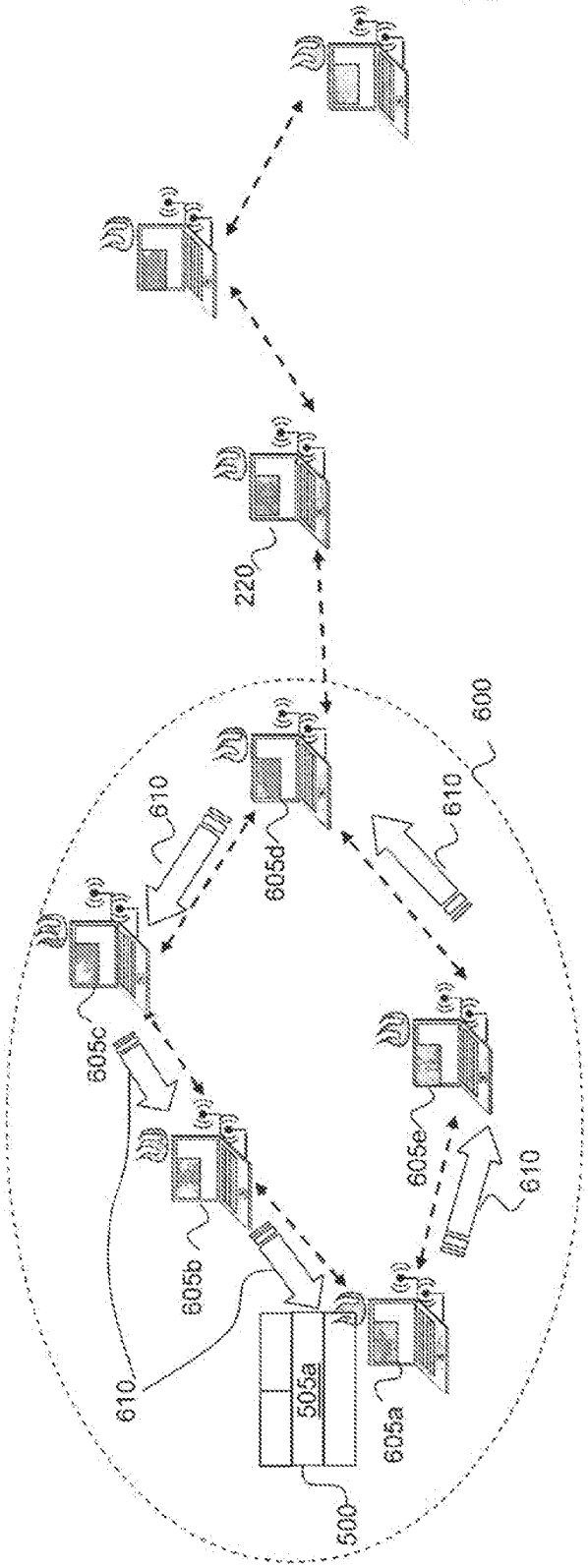


FIG. 6

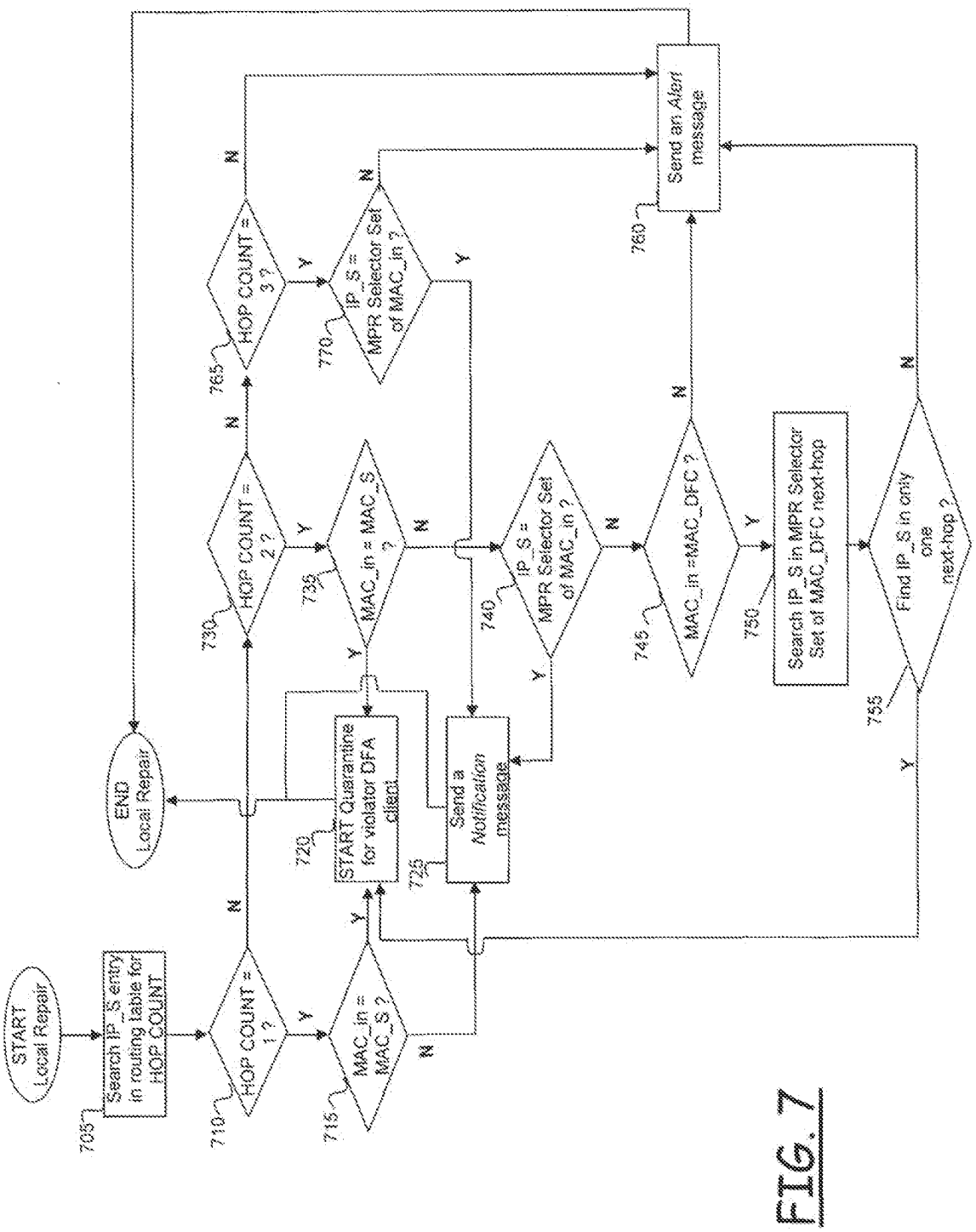

START Local Repair
Search IP_S entry in routing table for HOP COUNT
705
HOP COUNT = 1 ?
710
MAC_in = MAC_S ?
715
START Quarantine for violator DFA client
720
Send a Notification message
725
HOP COUNT = 2 ?
730
MAC_in = MAC_S ?
735
IP_S = MPR Selector Set of MAC_in ?
740
MAC_in =MAC_DFC ?
745
Search IP_S in MPR Selector Set of MAC_DFC next-hop
750
Find IP_S in only one next-hop ?
755
Send an Alert message
760
HOP COUNT = 3 ?
765
IP_S = MPR Selector Set of MAC_in ?
770
END Local Repair
Y
N

FIG. 7

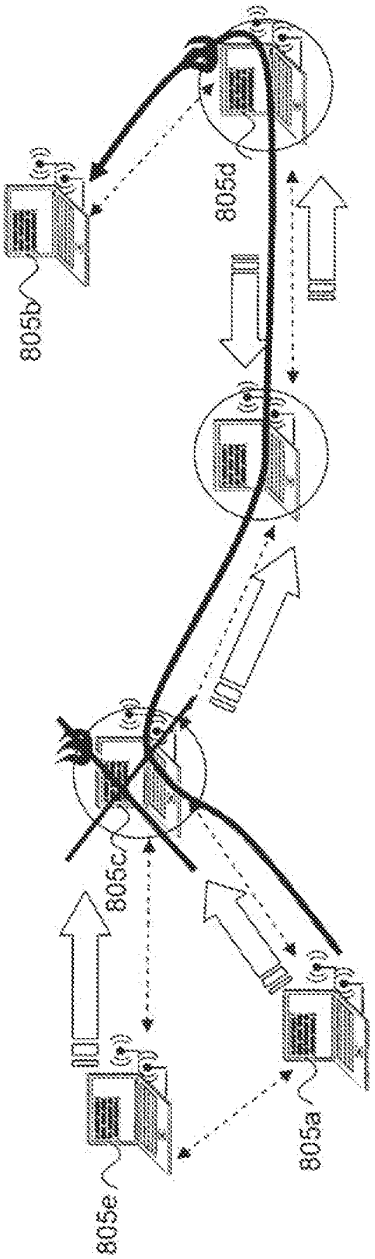


FIG. 8A

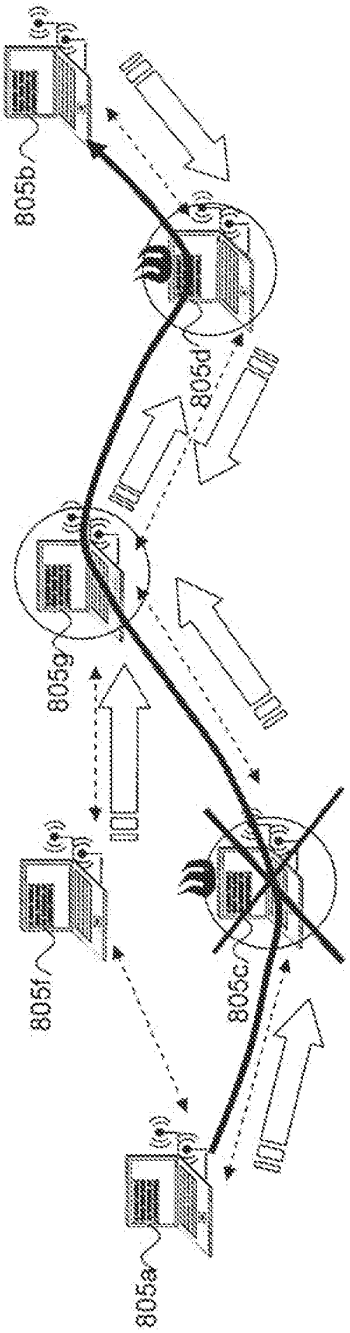


FIG. 8B

FIG. 8C

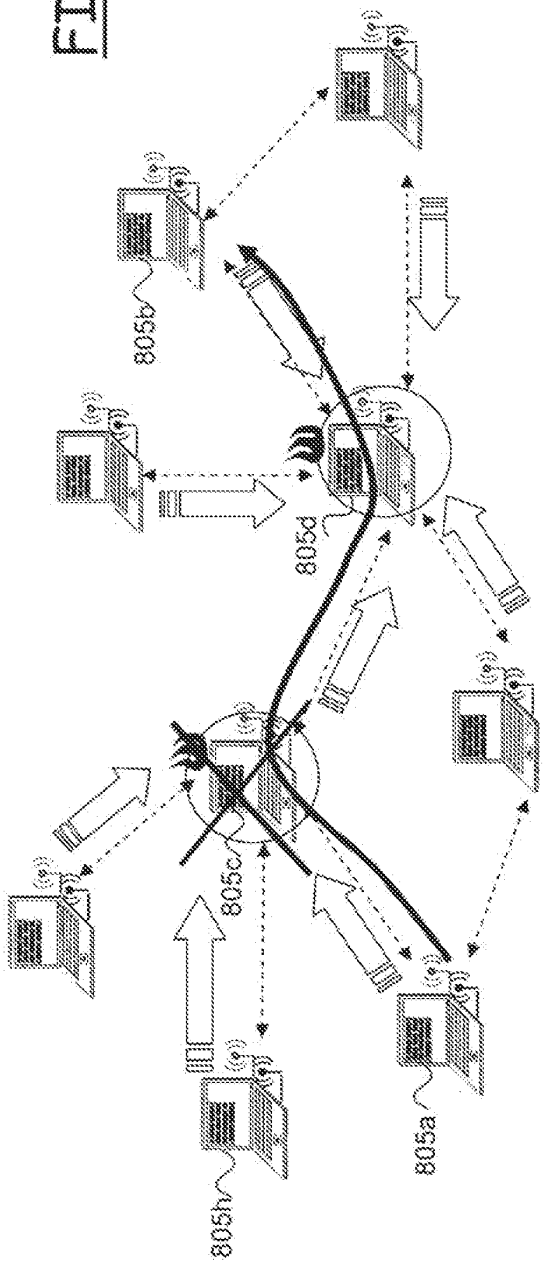
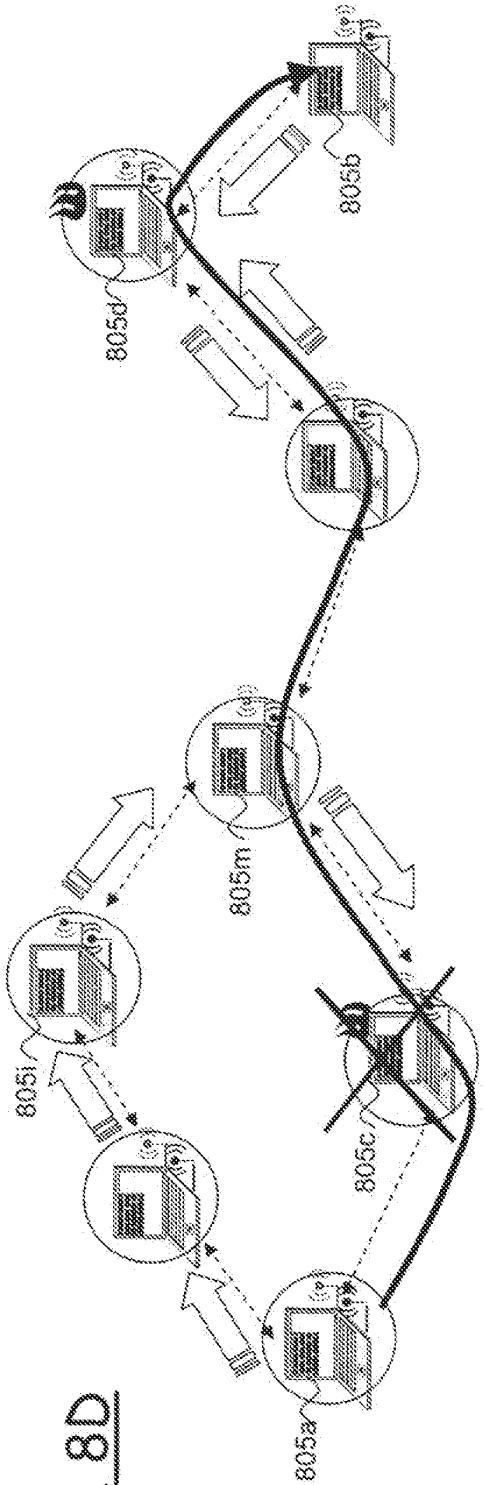


FIG. 8D



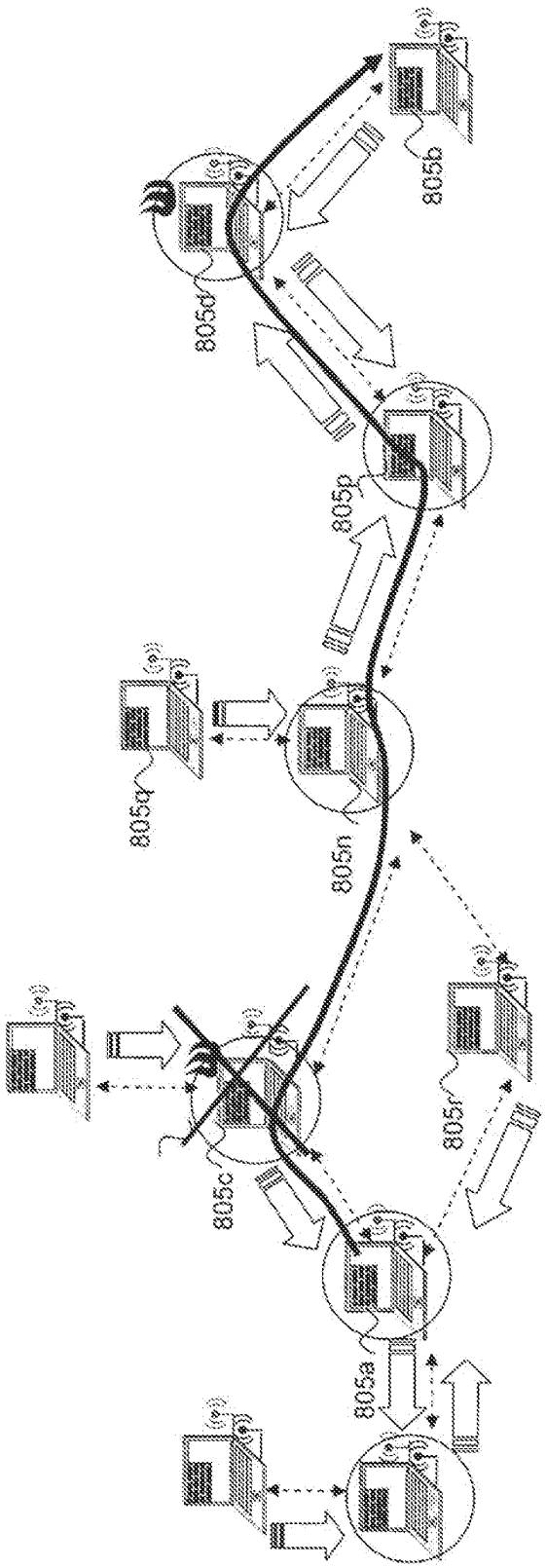


FIG. 8E

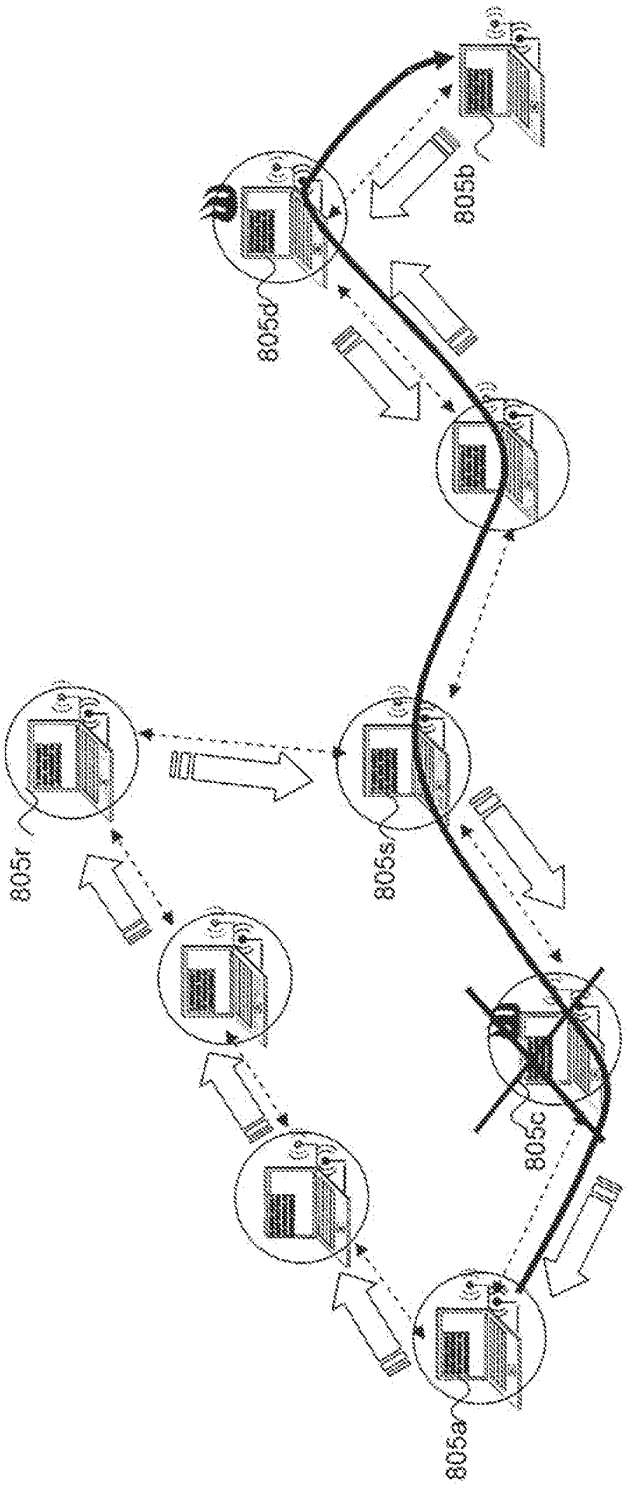


FIG. 8F

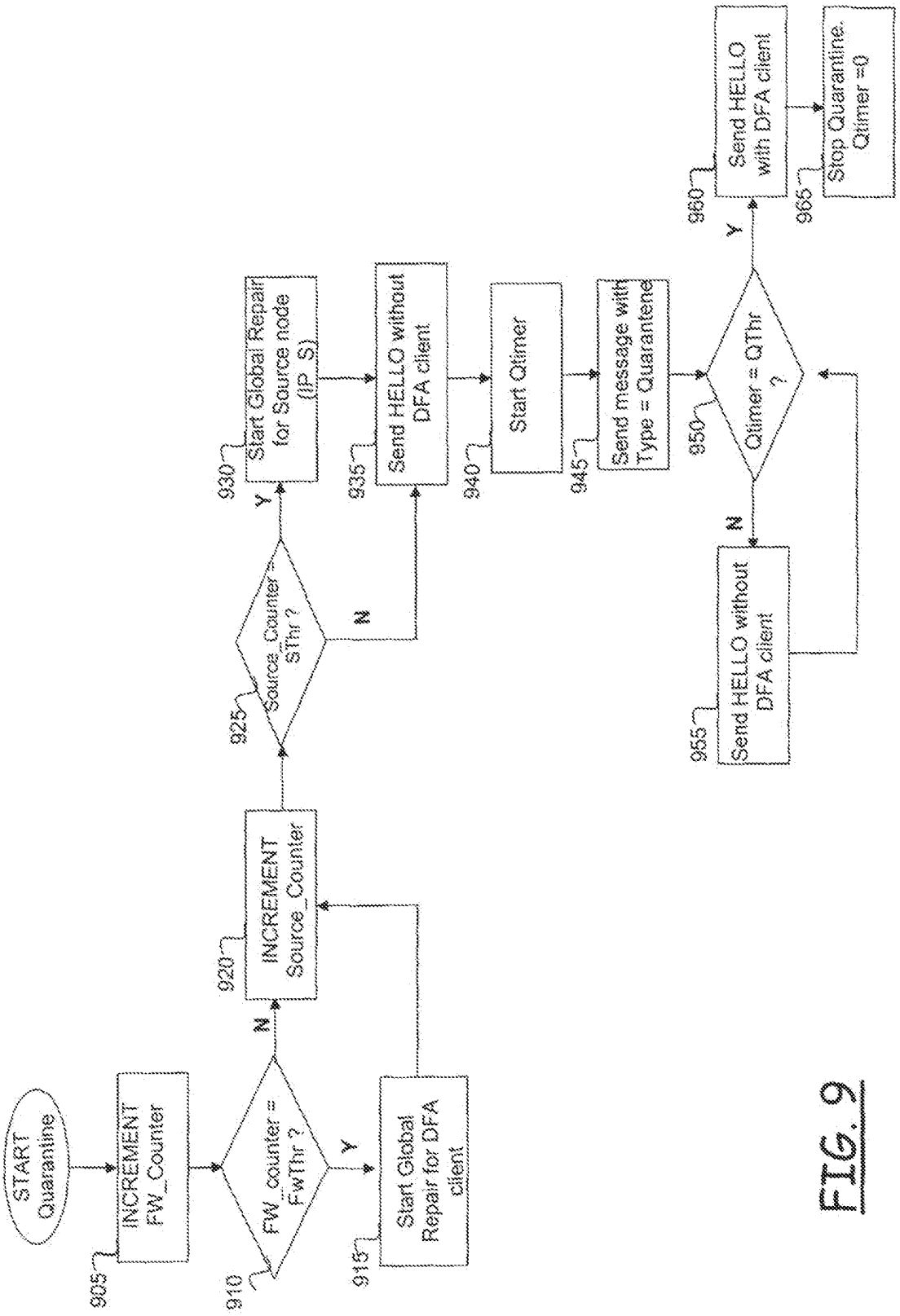

START Quarantine
INCREMENT FW_Counter
905
FW_counter = FwThr ?
910
Y
N
Start Global Repair for DFA client
915
INCREMENT Source_Counter
920
Source_Counter = SThr ?
925
Y
N
Start Global Repair for Source node (IP_S)
930
Send HELLO without DFA client
935
Start Qtimer
940
Send message with Type = Quarantene
945
Qtimer = QThr ?
950
Y
N
Send HELLO with DFA client
960
Stop Quarantine. Qtimer =0
965
Send HELLO without DFA client
955

FIG. 9

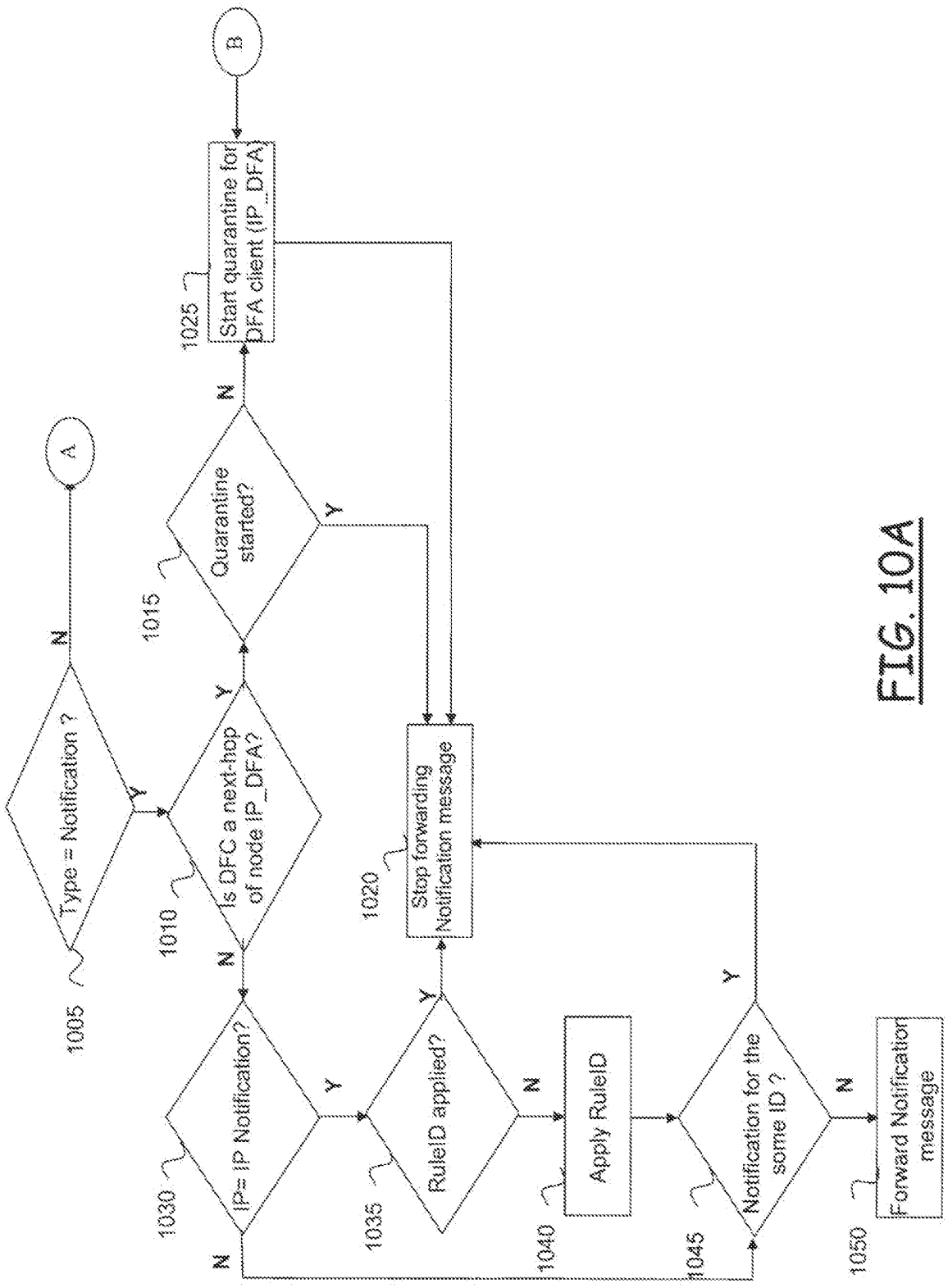

A
B
1005
Type = Notification ?
N
Y
1010
Is DFC a next-hop of node IP_DFA?
Y
N
1015
Quarantine started?
N
Y
1020
Stop forwarding Notification message
1025
Start quarantine for DFA client (IP_DFA)
1030
IP= IP Notification?
Y
N
1035
RuleID applied?
Y
N
1040
Apply RuleID
1045
Notification for the some ID ?
Y
N
1050
Forward Notification message

FIG. 10A

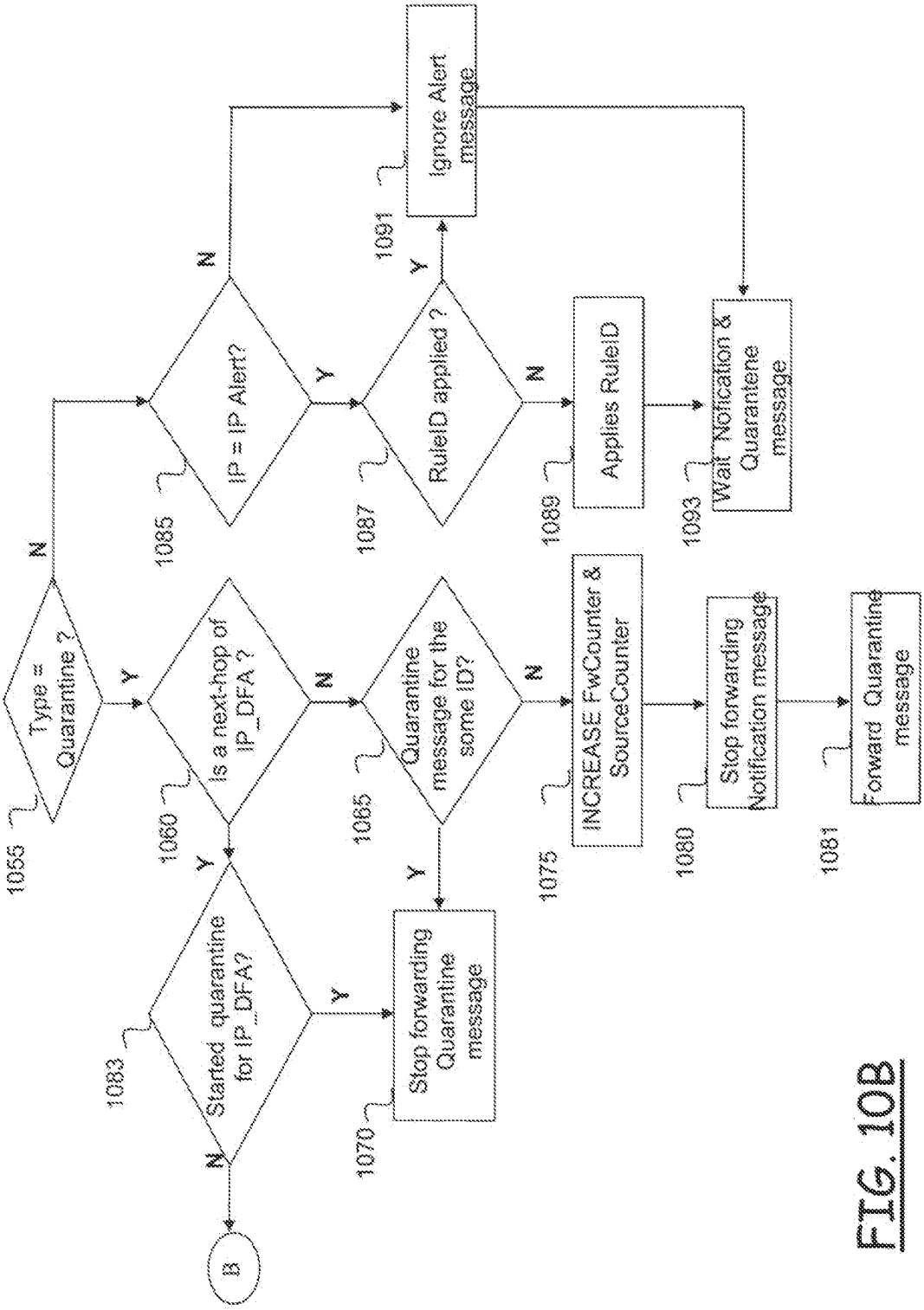

Type = Quarantine ?
1055
N
Y
Is a next-hop of IP_DFA ?
1060
N
Y
Started quarantine for IP_DFA?
1083
N
Y
B
IP = IP Alert?
1085
N
Y
RuleID applied ?
1087
N
Y
Applies RuleID
1089
Wait Notication & Quarantene message
1093
Ignore Alert message
1091
Quarantine message for the some ID?
1065
N
Y
INCREASE FwCounter & SourceCounter
1075
Stop forwarding Notification message
1080
Forward Quarantine message
1081
Stop forwarding Quarantine message
1070

FIG. 10B

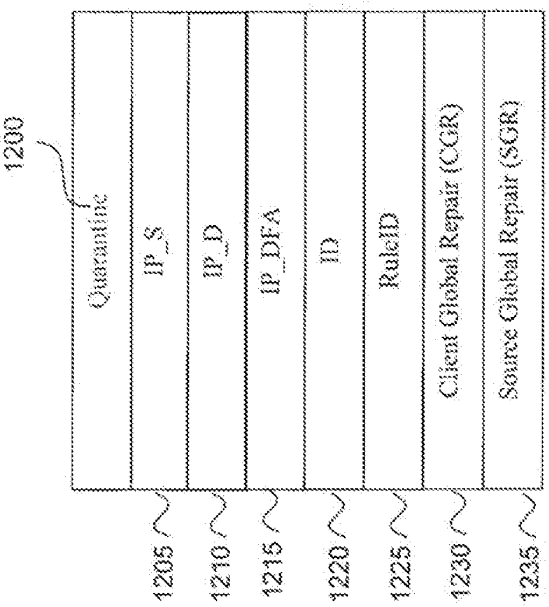


FIG. 12

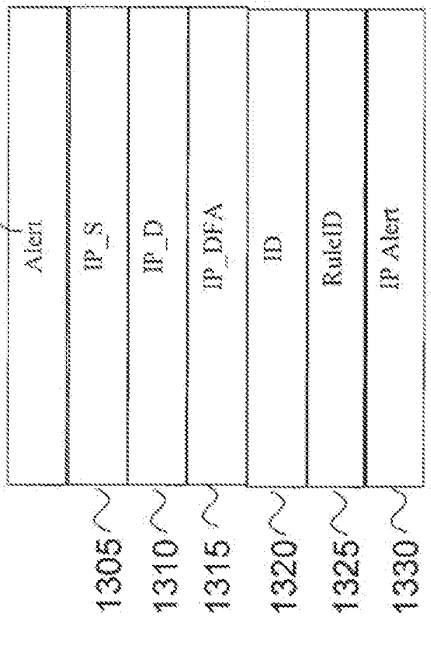


FIG. 13

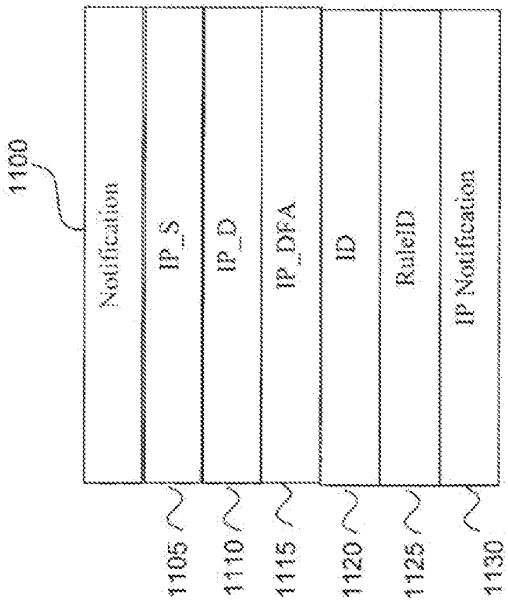


FIG. 11

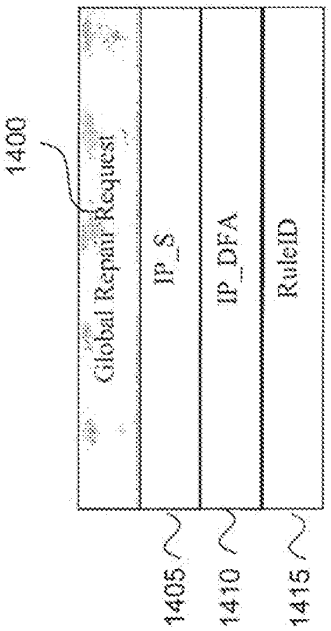


FIG. 14A

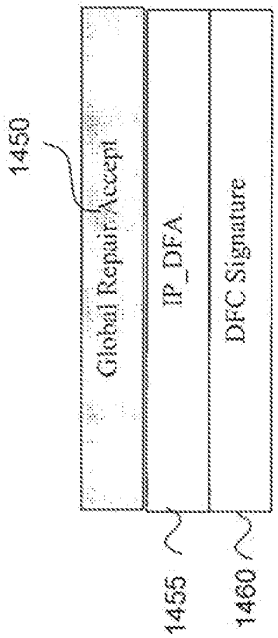


FIG. 14B

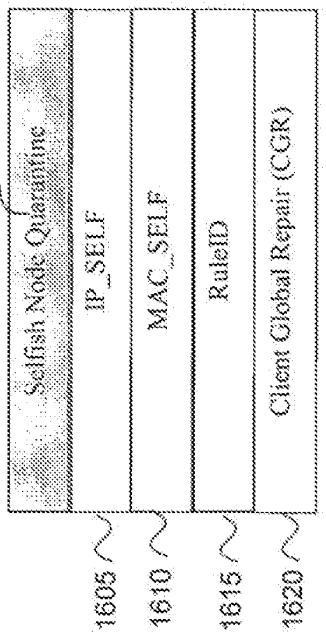
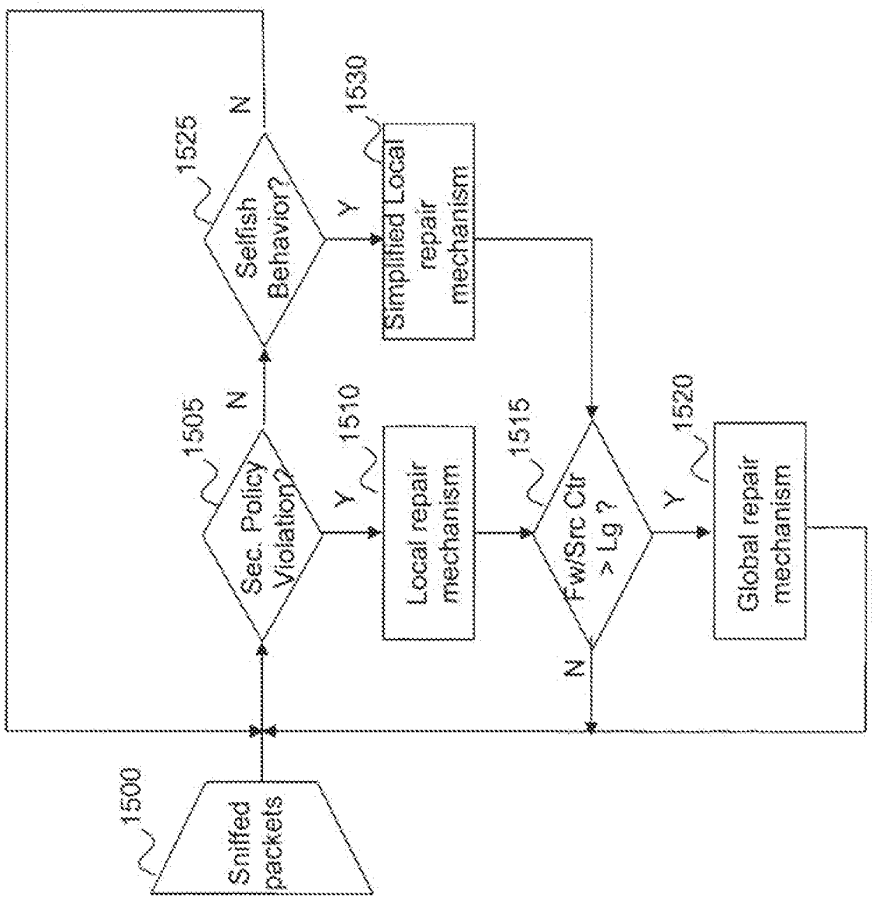


FIG. 16

FIG. 15


1500
Sniffed packets
1505
Sec. Policy Violation?
N
Y
1510
Local repair mechanism
1515
Fw/Src Ctr > Lg ?
N
Y
1520
Global repair mechanism
1525
Selfish Behavior?
N
Y
1530
Simplified Local repair mechanism

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/012606

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04L29/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DJORDJEVIC I ET AL: "An architecture for dynamic security perimeters of virtual collaborative networks" NETWORK OPERATIONS AND MANAGEMENT SYMPOSIUM, 2004. NOMS 2004. IEEE/IFIP SEOUL, KOREA APRIL 19-23, 2004, PISCATAWAY, NJ, USA,IEEE, 19 April 2004 (2004-04-19), pages 249-262, XP010712587 ISBN: 0-7803-8230-7	1-9, 13-23, 27,28
A	the whole document	10-12, 24-26

-/--

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 August 2007	29/08/2007

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Olaechea, Javier

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/012606

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RAMANUJAN R ET AL: "Intrusion-resistant ad hoc wireless networks" MILITARY COMMUNICATIONS CONFERENCE, MILCOM 2002. PROCEEDINGS. ANAHEIM, CA, OCT. 7 - 10, 2002, IEEE MILITARY COMMUNICATIONS CONFERENCE, NEW YORK, NY : IEEE, US, vol. VOL. 1 OF 2, 7 October 2002 (2002-10-07), pages 890-894, XP010631893 ISBN: 0-7803-7625-0	1-9, 13-23, 27,28
A	the whole document	10-12, 24-26
X	DJORDJEVIC I ET AL: "Architecture for secure work of dynamic distributed groups" CONSUMER COMMUNICATIONS AND NETWORKING CONFERENCE, 2004. CCNC 2004. FIRST IEEE LAS VEGAS, NV, USA 5-8 JAN. 2004, PISCATAWAY, NJ, USA,IEEE, 5 January 2004 (2004-01-05), pages 495-500, XP010696877 ISBN: 0-7803-8145-9	1-9, 13-23, 27,28
A	the whole document	10-12, 24-26
A	P.G. ARGYROUDIS,D.O'MAHONY: "Secure Routing for Mobile Ad hoc Networks"[Online] 23 August 2006 (2006-08-23), XP002447263 Retrieved from the Internet: URL:http://www.ctvr.ie/docs/EN_Pubs/secure-adhoc-routing.pdf> [retrieved on 2007-08-17] Sections 3, 4 and 5	1-28
A	CLAUSEN T ET AL: "Optimized Link State Routing Protocol (OLSR)" IETF STANDARD, INTERNET ENGINEERING TASK FORCE, IETF, CH, October 2003 (2003-10), XP015009408 ISSN: 0000-0003 cited in the application the whole document	1-28

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
14 February 2008 (14.02.2008)

(10) International Publication Number
WO 2008/017837 A1

(51) International Patent Classification:
H04M 1/2745 (2006.01) *H04M 1/725* (2006.01)

(21) International Application Number:
PCT/GB2007/003011

(22) International Filing Date: 8 August 2007 (08.08.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0615844.8 9 August 2006 (09.08.2006) GB

(71) Applicant *(for all designated States except US)*: **SYMBIAN SOFTWARE LIMITED** [GB/GB]; 2-6 Boundary Row, London SE1 8HP (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **SPOONER, David** [GB/GB]; Symbian Software Limited, 2-6 Boundary Row, London SE1 8HP (GB).

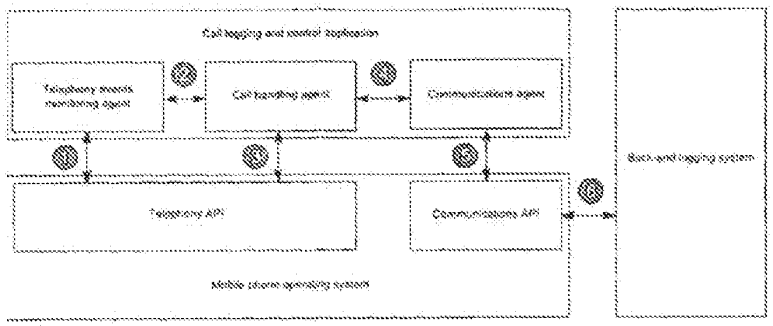
(74) Agent: KOEN, Sophia; Symbian Software Limited, 2-6 Boundary Row, London SE1 8HP (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) Title: MOBILE COMMUNICATIONS DEVICE WITH EVENT LOGGING



(57) Abstract: This invention relates to a mobile communications device programmed to provide logging of telephony activity. In one implementation, the programme provides near real-time notification and control of mobile phone calls. The notification is to a back-end logging system.

WO 2008/017837 A1

MOBILE COMMUNICATIONS DEVICE WITH EVENT LOGGING

FIELD OF THE INVENTION

This invention relates to a mobile communications device programmed to provide logging of phone call activity. In one implementation, the programme provides near real-time notification and control of mobile phone calls. The notification is to a back-end logging system.

Current mobile telephones typically include a log of incoming and outgoing phone calls, generally indicating which incoming calls were answered and which were missed.

Many mobile telephones also offer users a function to automatically divert all incoming calls to voicemail, without ringing.

Techniques are also known for barring outgoing phone calls to certain destinations, such as high-rate phone lines.

It is desirable to provide a phone call monitoring and logging system that provides enhanced control of call activity and enhanced logging capabilities, to enable functions to be automated and to improve a user's experience.

SUMMARY

In accordance with a first aspect of the present invention there is provided a mobile communications device programmed with: (a) a telephony events monitoring agent, registered with a telephony API provided by an operating system of the device so as to receive notifications of any telephony related events; (b) a call handling agent for processing telephony events received from the telephony events monitoring agent and arranged to communicate with the telephony API based on a configured call handling policy, the call handling agent being arranged to store logs of device telephony activity; (c) a communications agent that uses a communications API of the operating system to communicate the logs stored by the call handling agent to a back-end logging system using available communications means provided by the communications API.

In accordance with a second aspect of the invention there is provided a method of enabling a mobile communications device to log calls, comprising the steps of: (a) a telephony events monitoring agent registering with a telephony API provided by an operating system of the mobile device in order to receive notifications of any telephony related events; (b) a call handling agent processing telephony events received from the telephony events monitoring agent and arranged to communicate with the telephony API based on a configured call handling policy, the step of processing including storing logs of device call activity; (c) a communications agent using a communications API of the operating system to communicate the logs stored by the call handling agent to a back-end logging system using available communications means provided by the communications API.

In accordance with a third aspect of the invention there is provided a system for enabling mobile device calls to be logged, the system comprising multiple mobile communications devices as defined above and a back-end logging system arranged to receive logs from the mobile devices.

Preferred features of the invention are set out in the dependent claims.

DESCRIPTION OF DRAWINGS

The invention will now be described by way of example with reference to the accompanying drawing, in which:

Figure 1 shows a call handling arrangement including a back-end logging system.

DETAILED DESCRIPTION

Embodiments of the present invention enable a mobile communications device to intelligently control the handling of calls and telephony events, by reporting telephony events to a logging system and implementing a call handling policy using that logging system. A call handling agent receives input from a telephony events monitoring agent, and the call handling agent communicates with a telephony API of the device to enforce a policy for handling calls and telephony events.

In the preferred embodiment, a method is provided by which 1) telephony events on mobile devices (e.g. making and receiving calls, conferencing, missed calls) can be

reported to a back-end logging system (which also receives logs from landline telephones), and 2) telephony activities on the device can be controlled based on a dynamic policy (call barring for outgoing and incoming calls, call cost control, etc.).

The following series of steps explains how an embodiment of the invention can be used on a mobile telephone. The steps are illustrated in Figure 1.

1) The "Telephony events monitoring agent" registers with the mobile phone operating system's Telephony API in order to receive notifications of any telephony related events:
    a. Outgoing call;
    b. Incoming call;
    c. Call status change – dialling, answering, connecting, ringing, connected, hanging up, hold, etc.
2) The "Telephony events monitoring agent" notifies the "Call handling agent" of telephony events received from the Telephony API.
3) The "Call handling agent" processes messaging events received from the "Telephony events monitoring agent" and responds to the Telephony API based on configured policy. Call handling policy may cover the following functions:
    a. Log outgoing phone call and content;
    b. Log incoming phone call and content;
    c. Automatically answer incoming calls;
    d. Automatically disconnect incoming calls;
    e. Automatically disconnect outgoing calls;
    f. Determine response based on dynamic properties including message content or properties, date, time, location (GSM network location and/or GPS), phone call origin or destination, associated contact information (in business contacts group or personal contacts group, etc.), diary information, selected profile, etc.;
    g. Capture additional state information such as network information (network name, GSM country, area and cell, signal level), device status (memory consumption, battery level, selected profile);
    h. Determine schedule for communicating log information to the "Back-end logging system" (after each call event, every specific amount of time, at

specific times, once a specific amount of log information has been captured, etc.);

i. Invoke alternative communications client on device (e.g. automatically disconnect outgoing phone call and invoke voice-over-IP client to initiate call).

4) The "call handling agent" communicates call logging information to the "Communications agent" based on scheduling policy. Information may be augmented with additional device information such as device identification (IMEI, IMSI), network information (network type and identification, country, area and cell information). The "Communications agent" notifies the "Call handling agent" of any call handling policy modification requests received from the "Back-end logging system".

5) The "Communications agent" uses the mobile phone operating system's communications API to communicate with the "Back-end logging system" using available communications means provided by the communications API:

   a. A TCP/IP communication protocol (HTTP, FTP, SMTP, etc.)

   b. A GSM communication channel such as SMS, MMS.

6) The "Back-end logging system" may control the "Message handling agent" policy by communicating policy modifications using available communications means via the "Communications agent".

The back-end logging system is preferably provided remotely from the device. In this arrangement, the relatively small quantity of storage memory available on current mobile phones need not be impacted by the logged telephony data. Large quantities of storage memory can be provided on remote servers without reducing the available storage on the mobile device. However, it is possible to implement an embodiment of the invention in which the back-end logging system is provided in the device itself. Such an arrangement would have certain advantages: for example, the time taken for data or instructions to be transferred between the logging system and the communications agent could be much shorter.

The communication between the communications agent on the mobile telephone and the logging system preferably involves delivering all recent events logged by the call handling agent to the logging system, and then deleting them from the call handling

agent's local log. Once received at the back-end logging system, the recent events are compared with those already stored at the logging system, and updates are made accordingly.

The call handling policy defines the manner in which calls and telephony events are dealt with in the mobile device. It includes a series of rules which specify what should be done in particular circumstances or in response to a particular event. It can be used to control the content of the logging system, the disconnection of certain phone calls, the transition of a call between different states, or any other activity relating to telephony on the device.

In the case where the back-end logging system is used to direct the call handling policy, it could feed back to either the telephony API or the communications agent. It could formulate and modify policy based on criteria such as the amount of storage space available for storing new telephony information, the communication networks available to the device or the battery level. It could also be capable of receiving input from a user in order that the user can influence policy decisions. Alternatively it may be preferred that the user cannot access the back-end logging system, since making it directly accessible from the application level could represent a security risk – malicious or malfunctioning applications could potentially obtain or modify information from the logging system, thereby compromising the user's personal call information.

In the preferred embodiment the call handling policy is enforced at the call handling agent, which receives directions from the communications agent. The call handling agent communicates directly with the Telephony API in order to implement aspects of the call handling policy; however, the call handling agent could alternatively communicate with the API indirectly, via the telephony events monitoring agent or another element. In one embodiment, a user of the device can define at least some aspects of the call handling policy by means of interaction with the call handling agent, either directly or indirectly.

The scheduling policy defines the manner in which data is passed to the back-end logging system. It can define the type of information provided to the back-end logging system, as well as the timing of transmissions to the logging system. The scheduling policy could be determined on a dynamic basis according to the usage of the device's telephony applications. For example, when the device is being used frequently for calls – perhaps

during a weekday while the user is away from his office – the communication to the back-end logging system could be performed relatively frequently, whereas during a period of low usage – perhaps during a weekday while the user is at his office desk and primarily using his office phone and computer to send and receive communications – the communication to the back-end logging system could be performed infrequently. This variation in scheduling policy could be achieved by the use of a usage monitoring or prediction agent arranged to determine the current or likely telephony behaviour of the device.

In the preferred embodiment, all of the agents described above – the telephony events monitoring agent, the call handling agent, the communications agent, and the usage monitoring agent – are provided as a single call logging and control application. Alternatively, the logging and control aspects could be provided by separate applications, or some or all of the functionality could be provided within the operating system.

A simple example of the operation of the call logging and control arrangement of the present invention will now be given.

A user of a mobile telephone makes an outgoing call to an overseas destination. The telephony events monitoring agent receives a notification from the Telephony API indicating that a call has been made, and indicating the time of the call and the destination phone number. The call handling agent is informed of this call handling event by the telephony events monitoring agent. The call handling agent stores the outgoing call as an event in a store of recent events: an entry is made in the store to include the nature of the event (outgoing phone call), the time of the call and the number to which it is directed.

At this time, the call handling agent does not communicate back to the Telephony API since no action needs to be taken by the Telephony API in response to this event.

The mobile telephone in the example is currently configured to alert a user when the duration of an outgoing overseas or premium rate call has exceeded ten minutes, in order to warn the user of potentially expensive calls. This configuration has been selected by the user, and could be switched off or altered if preferred. This policy setting is invoked

by means of the back-end logging system, with which the user can communicate by password access to change his call handling policy.

In accordance with this policy setting, the call handling agent begins a timer when the new outgoing call is connected. After ten minutes, the call handling agent triggers the Telephony API to cause an alert to be sent to the user in the form of an audible notification. The user can then choose whether to terminate the call and avoid excessive costs.

The scheduling policy in force for the mobile telephone is fixed, and requires that all telephony events stored by the call handling agent are transmitted to the back-end logging system at midnight every day. Thus, at midnight on the day the user made his overseas call, the communications agent retrieves the new entry in the call handling agent's store and sends it over a GSM network to a remote back-end logging system.

It can be seen from the above description that the present invention provides a flexible technique for controlling and logging telephony activity on a mobile device. It will be understood by the skilled person that alternative implementations are possible, and that various modifications of the methods and implementations described above may be made within the scope of the invention, as defined by the appended claims.

CLAIMS

1. A mobile communications device programmed with:

(a) a telephony events monitoring agent, registered with a telephony API provided by an operating system of the device so as to receive notifications of any telephony related events;

(b) a call handling agent for processing telephony events received from the telephony events monitoring agent and arranged to communicate with the telephony API based on a configured call handling policy, the call handling agent being arranged to store logs of device telephony activity;

(c) a communications agent that uses a communications API of the operating system to communicate the logs stored by the call handling agent to a back-end logging system using available communications means provided by the communications API.

2. The device of Claim 1 in which the telephony related events include one or more of:

a) Making an outgoing call;

b) Receiving an incoming call;

c) Call status change, where the statues include dialling, answering, connecting, ringing, connected, hanging up, hold, missing a call.

3. The device of Claim 1 or 2 in which the configured policy includes one or more of the following:

a) Log outgoing call information;

b) Log incoming call information;

c) Automatically answer incoming calls;

d) Automatically disconnect incoming calls;

e) Automatically disconnect outgoing calls;

f) Determine response based on dynamic properties including date, time, location, call origin or destination, associated contact information, diary information, or selected profile;

g) Capture additional state information such as network information or device status;

h) Determine schedule for communicating log information to the back-end logging system;

i) Invoke alternative communications client on device.

4. The device of any preceding Claim in which the call handling agent communicates call logging information to the communications agent based on a scheduling policy.

5. The device of Claim 4 in which the call logging information may be augmented with additional device information such as device identification or network information.

6. The device of any preceding Claim in which the communications agent notifies the call handling agent of any call handling policy modification requests received from the back-end logging system.

7. The device of any preceding Claim in which the available communications means provided by the communications API includes one of:

a) A TCP/IP communication protocol (HTTP, FTP, SMTP, etc.)

b) A GSM communication channel such as SMS, MMS.

8. The device of Claim 1, arranged such that the back-end logging system can dynamically control the call handling policy by communicating policy modifications to the call handling agent using available communications means via the communications agent.

9. A method of enabling a mobile communications device to log calls, comprising the steps of:

(a) a telephony events monitoring agent registering with a telephony API provided by an operating system of the mobile device in order to receive notifications of any telephony related events;

(b) a call handling agent processing telephony events received from the telephony events monitoring agent and arranged to communicate with the telephony API based on a configured call handling policy, the step of processing including storing logs of device call activity;

(c) a communications agent using a communications API of the operating system to communicate the logs stored by the call handling agent to a back-end logging system using available communications means provided by the communications API.

10. A system for enabling mobile device calls to be logged, the system comprising multiple mobile communications devices as claimed in Claim 1 and a back-end logging system arranged to receive logs from the mobile devices.

11. The system of Claim 10 in which the back-end logging system is also arranged to log telephone calls received from mobile telephone and from landline calls.

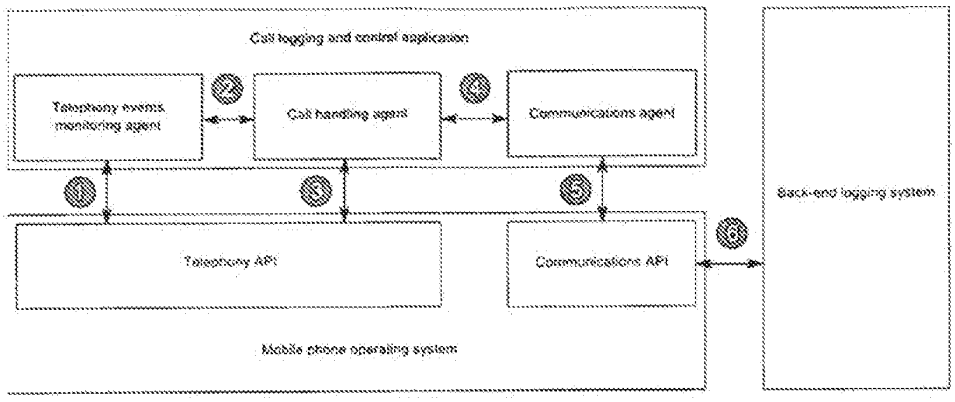

Call logging and control application
Telephony events monitoring agent
Call handling agent
Communications agent
Back-end logging system
Telephony API
Communications API
Mobile phone operating system

Figure 1

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2007/003011

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04M1/2745 H04M1/725

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04M H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/130631 A1 (MAGUIRE MICHAEL [CA] ET AL) 16 June 2005 (2005-06-16) abstract compounds 4-7 paragraph [0002] - paragraph [0004] paragraph [0011] - paragraph [0014] paragraph [0025] - paragraph [0034] paragraph [0037] - paragraph [0039] -/--	1-11

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 November 2007	14/11/2007

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kim-Mayser, Michael

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2007/003011

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/075097 A1 (LEHIKOINEN JAAKKO [FI] ET AL LEHIKOINEN JAAKKO [FI] ET AL) 7 April 2005 (2005-04-07) abstract figures 1-3c paragraph [0005] paragraph [0007] - paragraph [0010] paragraph [0016] - paragraph [0030] paragraph [0032] paragraph [0035]	1-11
A	US 2005/261003 A1 (FIEDLER RENO [US] ET AL) 24 November 2005 (2005-11-24) the whole document	1-11
A	EP 1 148 423 A (MITSUBISHI ELECTRIC CORP [JP]) 24 October 2001 (2001-10-24) the whole document	1-11
A	WO 02/078315 A (T MOBILE DEUTSCHLAND GMBH [DE]; VOEHRINGER GERRIT [DE]) 3 October 2002 (2002-10-03) the whole document	1-11

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2007/003011

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 2005130631	A1	16-06-2005	NONE		
US 2005075097	A1	07-04-2005	CN	1868202 A	22-11-2006
			EP	1671474 A2	21-06-2006
			WO	2005033828 A2	14-04-2005
			KR	20060085932 A	28-07-2006
US 2005261003	A1	24-11-2005	NONE		
EP 1148423	A	24-10-2001	CN	1335964 A	13-02-2002
			DE	60024230 D1	29-12-2005
			DE	60024230 T2	10-08-2006
			WO	0135229 A1	17-05-2001
			JP	3883869 B2	21-02-2007
			US	2001042103 A1	15-11-2001
WO 02078315	A	03-10-2002	AT	341897 T	15-10-2006
			CZ	20032905 A3	14-01-2004
			DE	10114964 A1	17-10-2002
			EP	1374551 A2	02-01-2004
			ES	2274031 T3	16-05-2007
			PL	362681 A1	02-11-2004
			US	2004116136 A1	17-06-2004

Form PCT/ISA/210 (patent family annex) (April 2005)

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2008 年8 月21 日 (21.08.2008)

PCT

(10) 国際公開番号
WO 2008/099802 A1

(51) 国際特許分類:
H04L 12/56 (2006.01)

(21) 国際出願番号: PCT/JP2008/052236

(22) 国際出願日: 2008 年2 月12 日 (12.02.2008)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願2007-031508 2007 年2 月13 日 (13.02.2007) JP

(71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 粟野 潤 (AWANO, Jun) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP).

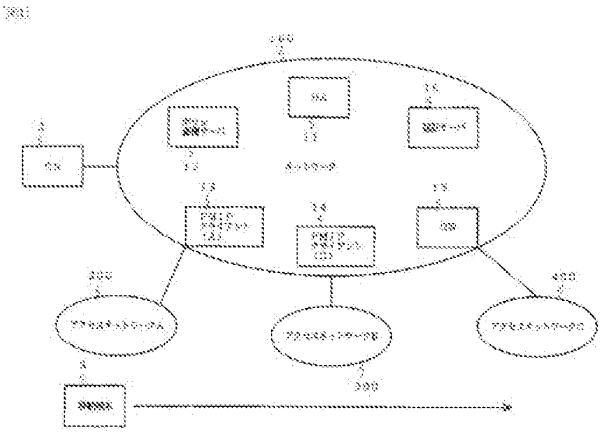
(74) 代理人: 宇高 克己 (UDAKA, Katsuki); 〒1010025 東京都千代田区神田佐久間町１－１４第二東ビル５階 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

/続葉有/

(54) Title: MOBILE TERMINAL MANAGEMENT SYSTEM, NETWORK DEVICE, AND MOBILE TERMINAL OPERATION CONTROL METHOD USED FOR THEM

(54) 発明の名称: 移動端末管理システム、ネットワーク機器及びそれらに用いる移動端末動作制御方法



100 NETWORK
12 POLICY ACCUMULATION SERVER
16 AUTHENTICATION SERVER
13 PMIP CLIENT (A)
14 PMIP CLIENT (B)
200 ACCESS NETWORK A
300 ACCESS NETWORK B
400 ACCESS NETWORK C
3 MOBILE TERMINAL

(57) Abstract: Provided is a mobile terminal management system which can rapidly decide whether a mobile terminal is to execute a PMIP operation or a CMIP operation in accordance with a network administrator.A PMIP client (A) and a PMIP client (B) have: a function to decide whether the mobile terminal is to execute the PMIP operation or the CMIP operation in accordance with terminal capability request information transmitted from a mobile terminal and a policy acquired from a policy accumulation server; and a function to transmit terminal operation control information which can control operation of the mobile terminal and IP setting information required upon decision whether the mobile terminal is to execute the PMIP operation or the CMIP operation. The mobile terminal has a function to transmit to the network, a report containing a use request of the CMIP function in accordance with the terminal side policy such as the user setting and service contents upon connection to the access network or upon a handover.

(57) 要約: ネットワークの管理者のポリシに応じて移動端末がＰＭＩＰ動作するか、ＣＭＩＰ動作するかを高速に決定可能な移動端末管理システムを提供する。ＰＭＩＰクライアント（Ａ）及びＰＭＩＰク

/続葉有/

WO 2008/099802 A1

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

ライアント（Ｂ）は、移動端末から送られた端末能力要求情報と、ポリシ蓄積サーバから取得したポリシに基づいて、移動端末をＰＭＩＰ動作とするか、ＣＭＩＰ動作とするかを決定する機能と、移動端末の動作を制御可能な端末動作制御情報と、移動端末がＰＭＩＰ動作、ＣＭＩＰ動作の何れとするかを決定した際に必要となるＩＰ設定情報とを含む通知を移動端末に送信する機能とを有する。移動端末は、アクセスネットワークに接続した際、あるいはハンドオーバした際等に、ユーザの設定やサービス内容等、端末側のポリシにしたがってＣＭＩＰ機能の使用要求を格納した通知をネットワークに送信する機能を有する。

明 細 書

移動端末管理システム、ネットワーク機器及びそれらに用いる移動端末動作制御方法

技術分野

[0001] 本発明は移動端末管理システム、ネットワーク機器及びそれらに用いる移動端末動作制御方法に関し、特にPMIP[Proxy MIP(Mobile Internet Protocol)]クライアントを配備したネットワークにおける移動端末の動作制御に関する。

背景技術

[0002] PMIPクライアントを配備したネットワークに、CMIP(Client MIP)機能[MIPv4(Mobile Internet Protocol version 4)機能]を備えた移動端末が接続した場合の動作例は、以下の通りである。

[0003] すなわち、移動端末がPMIPクライアントを有するネットワークに接続した場合には、MIPv4[RFC(Request For Comments)3344]に定義されたAA(Agent Advertisemet)を移動端末に対して送信し、移動端末がCMIP機能を備えているかどうかを確認する方法が考慮されている(例えば、非特許文献1参照)。

[0004] つまり、AAをCMIP端末が受信した場合には、それに反応してRRQ(Regisration Request:位置登録要求)を送信するが、このRRQが送信された場合、ネットワーク側は移動端末にCMIP機能があると判断し、移動端末はCMIPで動作する。

[0005] 一方、Simple IP(Internet Protocol)端末は、AAを理解できず、一般のIP端末がネットワークに接続した際と同様に、DHCP(Dynamic Host Configuration Protocol)によってアドレス取得を試みる。PMIPクライアントは、このDHCP requestを契機としてSimple IP端末の代わりにHA(Home Agent:ホームエージェント)に対して位置登録要求を送信する。

[0006] 上記の処理動作によれば、CMIP端末は固定的にCMIP動作することになる。但し、ネットワークを管理するオペレータによっては、CMIP端末に対してもPMIPで動作させたい場合がある。

[0007] しかしながら、現在、PMIP機能を有するネットワークにCMIP端末が接続した時に

は、移動端末側の要求やネットワークでのポリシを鑑みつつ、移動端末の動作(PMIP動作とするか、CMIP動作とするか)を行う仕組みが存在しない。

[0008] 非特許文献1:"WiMAX End-to-End Network Systems Architecture(Stage 2:Architecture Tenets, Reference Model and Reference Points)[Part2], 7.8 CSN Anchored Mobility Management"

非特許文献2:"Mobility Management using Proxy Mobile IPv4", http://www.ietf.org/internet-drafts/draft-leung-mip4-proxy-mode-01.txt

非特許文献3:"Proxy Mobile IPv6", http://www.ietf.org/internet-drafts/draft-sgundave-mipv6-proxymipv6-00.txt

非特許文献4:"Network Based Layer 3 Connectivity and Mobility Management for IPv6", http://www.ietf.org/internet-drafts/draft-chowdhury-netmip6-01.txt

発明の開示

発明が解決しようとする課題

[0009] PMIPクライアントを配備したネットワークに、MIP機能(MIPv4機能)を備えた移動端末が接続した場合の動作においては、移動端末に対して送信したAAに反応して移動端末が位置登録要求(RRQ)を送信してくるか、AAを無視してDHCPによってアドレス取得にくるかに応じて移動端末におけるMIP機能の有無を判定し、MIP機能がなければPMIPクライアントが移動端末に代わってRRQを送信し、MIP機能があればPMIPクライアントが移動端末の代理を行わず、移動端末自身がCMIP動作によってRRQを送信している。

[0010] 上記の処理動作によれば、CMIP端末は固定的にCMIP動作することになる。ネットワークを管理するオペレータによっては、CMIP端末に対してもPMIPで動作させたい場合があるが、CMIP端末がMIP機能を使わず、PMIP動作とするか、MIP機能を使ってCMIP動作とするかをネットワーク側から自由に制御する仕組みが存在しない。

[0011] そこで、本発明の目的は上記の問題点を解消し、ネットワークの管理者のポリシにし

たがって移動端末がPMIP動作するか、CMIP動作するかを高速に決定することができる移動端末管理システム、ネットワーク機器及びそれらに用いる移動端末動作制御方法を提供することにある。

課題を解決するための手段

[0012] 本発明による移動端末管理システムは、PMIP[Proxy MIP(Mobile Internet Protocol)]機能を持つネットワーク機器が配備されたネットワークにおいて移動端末の動作を制御する移動端末管理システムであって、 前記ネットワーク側から前記移動端末側に前記PMIP機能を提供可能か否かを通知している。

[0013] 本発明によるネットワーク機器は、PMIP[Proxy MIP(Mobile Internet Protocol)]機能を持つネットワーク機器が配備されたネットワークにおいて移動端末の動作を制御する移動端末管理システムを構成するネットワーク機器であって、

前記移動端末側に対して前記PMIP機能を提供可能か否かを通知している。

[0014] 本発明による移動端末動作制御方法は、PMIP[Proxy MIP(Mobile Internet Protocol)]機能を持つネットワーク機器が配備されたネットワークにおいて移動端末の動作を制御する移動端末管理システムに用いる移動端末動作制御方法であって、

前記ネットワーク側から前記移動端末側に前記PMIP機能を提供可能か否かを通知している。

発明の効果

[0015] 本発明は、上記のような構成及び動作とすることで、ネットワークの管理者のポリシにしたがって移動端末がPMIP動作するか, CMIP動作するかを高速に決定することができるという効果が得られる。

図面の簡単な説明

[0016] [図1]本発明の第1の実施の形態による移動端末管理システムの構成を示すブロック図である。

[図2]NVSEのフォーマットを示す図である。

[図3]図1の移動端末の構成例を示すブロック図である。

[図4]図1のPMIPクライアント(A)13の構成例を示すブロック図である。

[図5]図1のGWの構成例を示すブロック図である。

[図6]本発明の第1の実施の形態におけるNVSEが格納されたMIPv4信号を受信したMIPv4ノードの処理を示すフローチャートである。

[図7]本発明の第1の実施の形態において移動端末3がネットワークのポリシによってPMIP動作とする場合の動作を示すシーケンスチャートである。

[図8]本発明の第1の実施の形態において移動端末3からの要求とネットワークのポリシとによって移動端末3の動作を決定する手順を示すフローチャートである。

[図9]本発明の第1の実施の形態において移動端末3がアクセスネットワーク(A)200に接続した時にネットワークのポリシによってCMIP動作が許可される場合の動作を示すシーケンスチャートである。

[図10]本発明の第1の実施の形態において移動端末3がアクセスネットワーク(C)400に接続した場合の動作を示すシーケンスチャートである。

[図11]本発明の第2の実施の形態による移動端末管理システムの動作を示すシーケンスチャートである。

[図12]本発明の第2の実施の形態において移動端末3がアクセスネットワーク(A)200に接続した場合にネットワークのポリシによってCMIP動作が許可された場合の動作を示すシーケンスチャートである。

[図13]本発明の第2の実施の形態において移動端末3がアクセスネットワーク(C)400に接続した場合の動作を示すシーケンスチャートである。

符号の説明

[0017] 2 CN

3 移動端末

11 HA

12 ポリシ蓄積サーバ

13 PMIPクライアント(A)

14 PMIPクライアント(B)

15 GW

31 CMIP動作手段

32 Simple IP動作手段
33 接続・動作検出手段
34 ポリシ蓄積手段
35 端末動作要求送信手段
36 端末動作情報受信手段
37 動作決定手段
100 ネットワーク
131 PMIP動作手段
132 FA動作手段
133 端末動作要求受信手段
134 ポリシ通信手段
135 動作決定手段
136 認証処理手段
137 端末動作情報送信手段
151 GW動作手段
152 FA動作手段
153 端末動作要求受信手段
154 認証処理手段
155 端末動作情報送信手段
200 アクセスネットワーク(A)
300 アクセスネットワーク(B)
400 アクセスネットワーク(C)

発明を実施するための最良の形態

[0018] 本発明の移動端末管理システムは、移動端末がMIP(Mobile Internet Protocol)[以後、PMIP(Proxy MIP)に対してClient側のMIP機能であることを強調し、Client MIP(CMIP)とする]機能を有している場合(以下、この移動端末をCMIP端末とする)、CMIP端末がネットワークに接続した際、あるいはハンドオーバした際に、CMIP能力の通知及びCMIP機能の使用をネットワーク機器へ要求する機能を

、CMIP端末に備えている。

[0019] また、ネットワークでは、該ネットワークの管理者のポリシに基づいて、CMIP端末の動作を制御可能な情報、及びIP(Internet Protocol)設定情報を格納した通知をCMIP端末に対して行う。これにより、CMIP端末がPMIP端末(つまり、Simple IP端末)あるいはCMIP端末としての動作のいずれとするかを明確に決定することが可能となる。

[0020] さらに、PMIP端末として動作する場合、CMIP端末として動作する場合のいずれおいても、移動端末の各種IP情報[移動端末のHoA(Home Address)、CoA(Care-of Addres)、HA(Home Agent;ホームエージェント)アドレス]の設定を高速に完了することが可能となる。

[0021] そのため、本発明の移動端末管理システムでは、ネットワーク接続やハンドオーバ時において、サービス提供可能となるまでの時間を短縮することが可能となる。この時、端末能力通知、CMIP使用の要求、及びネットワークからの端末制御通知に、MIPv4(Mobile Internet Protocol version 4)の場合にはAS(Agent Solicitation)、AA(Agent Advertisement)、MIPv6(Mobile Internet Protocol version 6)の場合にはRS(Router Solicitation)、RA(Router Advertisement)といったMIPやIP端末が通常に使用するプロトコルを拡張して行うことで、本発明に対応した端末だけでなく、一般的なCMIP端末、あるいはSimple IP端末も同時に収容することが可能であり、現存するシステムへの本発明の適用も容易となる。

[0022] また、本発明の移動端末管理システムでは、端末とネットワークとの間の通知の数も削減可能となるために、通信帯域の効率的な使用が可能となる。移動通信では無線通信技術を使用するが、無線通信技術は、多くの場合、有線通信技術と比べて帯域的に劣るため、帯域の効率的な使用は特に重要となる。

[0023] このように、本発明の移動端末管理システムは、PMIP機能を備えたネットワークに、CMIP機能を備えたCMIP端末が接続した際に、CMIP端末のポリシによってCMIPでの動作を要求する通知を行う。一方、ネットワークがその通知によって取得したCMIP端末の要求とネットワーク側のポリシとによって、CMIP端末をCMIPで動作させるか、PMIPで動作させるかを制御可能な通知を行うと共に、通信に必要な情報(例

えば、IPアドレス情報)を通知することを特徴としている。

[0024] この結果、本発明の移動端末管理システムでは、CMIP端末がネットワークに接続した際、あるいはハンドオーバした際に、PMIP動作(ネットワーク側のPMIP機能を使用)とするか、CMIP動作(自らのCMIP機能を使用)とするかを決定することが可能となり、さらにIP設定を完了することが可能となるため、高速に通信可能な状態となり、電源投入後の快適なサービスの開始、あるいはハンドオーバ時のサービス途絶の回避ないしは途絶時間の短縮を可能とする。

[0025] 本発明の移動端末管理システムは、さらに以下の(1)～(3)の特徴を有している。

(1)本発明の移動端末管理システムでは、既存のMIPのプロトコル、あるいは一般に使用されるプロトコルを流用することで、既存のシステムへの変更を最小限とする。

(2)本発明の移動端末管理システムでは、本発明をサポートしたPMIPクライアントを配置したネットワークに、本発明をサポートしないCMIP端末やCMIP機能を備えない一般的なIP端末(以後、この端末をSimple IP端末とする)が接続した場合にも、それらの端末がネットワークに接続した際に発生する信号交換と共通化することが可能となり、一般的なIP端末及びCMIP端末との共存が可能であり、本発明を既存システム、あるいは既に検討が進んでいるシステムに適用することを容易とする。

(3)本発明をサポートしたネットワークに、本発明をサポートしたMIP端末が接続した場合には、ネットワーク側のポリシが許せば、MIP端末側の意図によって、PMIP動作とするか、CMIP動作とするかの動作モードを選択することが可能となる。これによって、端末側でハンドオーバタイミングを制御することが可能となる。また、端末が、サービスの種別によって、HA(Home Agent)を選択したり、複数のHAに位置登録をしたりすることが可能となり、ユーザに提供するサービスの幅を広げることが可能となる。

[0026] 尚、HAに相当するのは、3GPPにおいては、SAE－A(System Architecture Evolution－Anchor)である。3GPP2, WiMAX ForumではそのままHAと呼称される。また、PMIPクライアントに相当するのは、3GPPにおいては、3GPP－A(3GPP－Anchor)、MME(Mobility Management Entity)／UPE(User Plane Entity)、PDG(Packet Data Gateway)である。また、3GPP2では、PDS

N(Packet Data Serving Node)、PDIF(Packet Data Interworking Function)である。さらに、WiMAX ForumにおいてはASN(Access Service Network)－GWである。

[0027] 続いて、本発明の詳細について、以下に図面を参照して説明する。

[0028] 図1は本発明の第1の実施の形態による移動端末管理システムの構成を示すブロック図である。図1において、本発明の第1の実施の形態による移動端末管理システムは、ネットワーク100と、アクセスネットワーク(A)200と、アクセスネットワーク(B)300と、アクセスネットワーク(C)400と、CN(Correspondent Node)2と、移動端末3とから構成されている。

[0029] ネットワーク100内には、HA(Home Agent)11と、ポリシ蓄積サーバ12と、PMIP[Proxy MIP(Mobile Internet Protocol)]クライアント(A)13と、PMIPクライアント(B)14と、GW(GateWay)15とが設けられている。

[0030] ネットワーク100は、移動端末3に対して各種サービスを提供するプロバイダが管理するネットワークである。ネットワーク100は、例えば、3GPP(3rd Generation Partnership Project)においてCore Networkと呼ばれるネットワークに対応し、また、WiMAXにおいてはCSN(Connectivity Service Network)に対応するネットワークである。

[0031] HA11はRFC(Request For Comments)3344、あるいはRFC3775に定義されたHAであり、CMIP(Client MIP)端末、あるいはPMIPクライアントによる位置登録要求によって、移動端末のHome Address宛のパケットを、移動端末の移動先にカプセル化によって転送する機能を有している。

[0032] 本実施の形態でのHA11は、一般的なHAと同等の機能を有するものである。当該HAは、例えば、3GPPではSAE－A(System Architecture Evolution－Anchor)、3GPP－A(3GPP－Anchor)というように、MIPのHAの移動端末の移動先へパケットを転送する機能に加え、他の機能も一緒に実装されたり、異なる名称で呼ばれたりするが、MIPとしてのHA機能が実装されているノードは全て、HA11として示したノードの範疇に入る。

[0033] ポリシ蓄積サーバ12は、CMIP端末をSimple IP(Internet Protocol)端末とし

て動作させるか、CMIP端末として動作させるかといったネットワークのポリシが蓄積されたサーバであり、PMIPクライアント(A)13、PMIPクライアント(B)14からのポリシ要求によって、ネットワークのポリシを応答する。このポリシは、ネットワーク単位で共通のポリシとしても良いし、アクセスネットワーク単位でのポリシとしても良い。あるいは、ユーザ、あるいは移動端末単位で異なるポリシとしても良い。

[0034] アクセスネットワーク単位でポリシを適用する場合には、ポリシ要求に、PMIPクライアントあるいはGWの識別子を含める。また、ユーザ単位でポリシを適用する場合には、ポリシ要求に、RFC4282で規定されたユーザのNAI(Network Access Identifier)等のユーザ識別子を格納する。同様に、移動端末単位とする場合には、端末識別子をポリシ要求に格納する。

[0035] 一方、固定的に同一のポリシを適用する場合には、ポリシ蓄積サーバ12は省略することが可能である。この場合には、PMIPクライアント(A)13、PMIPクライアント(B)14、GW15各々の設定ファイルに、CMIP端末の動作ポリシを設定する。

[0036] ポリシ蓄積サーバ12に格納するポリシ情報は、例えば、CMIP端末にPMIPとして動作させることを強制する"PMIP強制"、CMIP端末にPMIP動作、あるいはCMIP動作の選択を許可する"CMIP許可"、CMIP端末にCMIP動作を強制する"CMIP強制"とすることができるが、これらに限らない。

[0037] PMIPクライアント(A)13及びPMIPクライアント(B)14は、一般的なPMIPクライアント機能に加え、移動端末3から送られた端末動作要求情報と、ポリシ蓄積サーバ12から取得したポリシに基づいて、移動端末3をPMIP動作とするか、CMIP動作とするかを決定する機能を備え、その後、移動端末3の動作を制御可能な端末動作制御情報と、移動端末3がPMIP動作、CMIP動作の何れとするかを決定した際に必要となるIP設定情報とを含む通知を移動端末3に送信する機能を有する。

[0038] ここで、一般的なPMIP機能とは、移動端末3に代ってCMIPの位置登録要求をHA11に送信する機能である。位置登録要求とは、MIPv4であればRRQ(Registration Request)であり、MIPv6であればBU(Binding Update)である。

[0039] PMIPクライアントの仕様は、MIPv4をベースにするか、MIPv6をベースにするか、あるいはIPv4、IPv6双方の端末を収容可能とするかどうか、位置登録要求を代理

送信する契機を何にするか、等いくつかの相違点を有する複数のドラフト(上記の非特許文献2～4)がIETF(Internet Engineering Task Force)に提出されている。

[0040] しかしながら、いずれにおいても、上述したように、端末の代理で位置登録要求をHA11に送信するPMIP機能を持つノードが規定されており、PMIP機能を有するノードが、本発明のPMIPクライアント(A)13, PMIPクライアント(B)14に相当する。今後新たなドラフトが提出される、あるいは正式な標準ができたとしてもPMIP機能を有するノードであれば、本発明の要素を盛り込めば、本実施の形態のPMIPクライアントに相当するノードとなる。

[0041] 本実施の形態では、上記のように、PMIPクライアント(A)13、PMIPクライアント(B)14は、一般的なPMIP機能と、本発明の移動端末3の動作制御情報を通知する機能とを共に備えたノードとしているが、これらの機能は別のノードとして実装されても構わない。

[0042] PMIPクライアント(A)13, PMIPクライアント(B)14はそれぞれ、ネットワーク100と、アクセスネットワーク(A)200、アクセスネットワーク(B)300との境界部に配置されるが、ユーザあるいは通信I／F(インタフェース)の認証を行うための認証機能を、さらにPMIPクライアント(A)13、PMIPクライアント(B)14に有する構成としても構わない。上記のような機能を含め、本発明のPMIPクライアントとしたノードの機能配備は、実施する状況によって様々なものが考えられ、本実施の形態で想定した機能配備には限らない。

[0043] GW15は、ネットワーク100とアクセスネットワーク(C)400との境界に配置されたルータであり、移動端末3がネットワーク100内のノードと通信するにあたって、ユーザ認証や、通信I／Fの認証を行うための認証機能を備えていても良い。上記のような一般的なルータ、あるいは認証GWとしての機能に加え、本発明のGW15は、移動端末3の動作がCMIPとなるような端末動作制御情報を移動端末3に送ると共に、移動端末3がCMIP動作をする際に必要となるIP設定情報を含む通知を移動端末3に送信する機能を有している。

[0044] PMIPクライアント(A)13、PMIPクライアント(B)14, GW15が、移動端末3に対し

て端末動作制御情報及びIP設定情報を通知する場合、本発明においては、MIPv4、MIPv6が一般的に使用するプロトコルを拡張して情報を格納する方法とする。

[0045] 具体的に説明すると、MIPv4の場合にはAA(Agent Advertisement)に、MIPv6の場合にはRA(Router Advertisemet)にそれぞれ情報を格納する。この時、必ずしも、AA、RAのような一般的に使用するプロトコルを使用する必要はないが、AA、RAを使用することで、(1)AAあるいはRAの他に新たな通知を送る必要がなく、通信帯域の効率的な使用が可能、(2)本発明をサポートしたCMIP端末以外にも、一般的なCMIP端末及びSimple IP端末も矛盾無く収容することが可能、(3)AA、RAに格納可能な一般的な情報要素に入れる情報を工夫することで、本発明をサポートしないCMIP端末でも、本発明をサポートするCMIP端末に対して効率や自由度の面で劣るものの、CMIP動作、PMIP動作の制御が可能、といったメリットが得られる。

[0046] 尚、MIPv4のAAあるいはRAを同一リンク内の複数の移動端末に送信する場合には、特定の端末のIP設定情報を含めることはできないが、本発明の好適な適用先として想定している3GPPや3GPP2、WiMAX Forumで検討されているネットワークは、PMIPクライアントやGWと、移動端末とが移動端末単位で個別のリンクで接続される。したがって、AA、RAは特定の移動端末に送信されることになり、特定の移動端末のIP設定情報を含むこととしても問題ない。

[0047] また、上記のような場合ではなく、一般的なネットワークに適用する場合においても、特定の移動端末によって送信されたASあるいはRSへの応答としてUnicastでAA、RAを送信する場合にもIP設定情報を含めることができる。いずれにせよ、本実施の形態ではAA、RAは特定の移動端末に送信される状況を想定している。但し、IP設定情報を含めず、移動端末の動作制御情報を格納する場合には、複数の移動端末にAAやRAが送信される場合にも適用することができる。

[0048] 認証サーバ160は、RADIUSやDIAMETER等の一般的な認証サーバであり、アクセスネットワーク(A)200、アクセスネットワーク(B)300、アクセスネットワーク(C)400に移動端末3が接続した場合には、それぞれ、PMIPクライアント(A)13、PMIPクライアント(B)14、GW15からの認証要求に応じて移動端末3のユーザ、あるいは

端末自身の認証処理を行い、結果を応答する機能を有している。

[0049] また、認証をパスした際には、MIPクライアント(A)13、PMIPクライアント(B)14、GW15は、移動端末3に割当てるべきHoA(Home Address)やHAアドレス、HAに位置登録要求を行う際に使用する鍵情報を認証サーバ16からダウンロードすることができる。これらの機能は、上記のRADIUSやDIAMETERによる一般的な認証サーバで実現できることであり、それらの詳細な説明については省略する。

[0050] この例では、HA11、ポリシ蓄積サーバ12、認証サーバ16は、ネットワーク100の中に配置されているが、別のネットワークに配置されたものを使用することもできる。例えば、ネットワーク100の管理者が行っているサービスの加入者の移動端末3が、別の管理者が管理するネットワークに接続する場合、HA11や認証サーバ16はネットワーク100内のものを使用して、ポリシ蓄積サーバは在圏のものが使用される場合が考えられる。

[0051] CN2は、移動端末3の通信相手となる端末、あるいはサーバである。移動端末3の通信相手となる端末の場合には、移動端末3と同等の機能を有する移動端末であっても良い。

[0052] 移動端末3は、一般的なMIPv4、あるいはMIPv6機能に加え、アクセスネットワークに接続した際、あるいはハンドオーバした際等に、ユーザの設定やサービス内容等、端末側のポリシにしたがってCMIP機能の使用要求を格納した通知をネットワークに送信する機能を備えている。

[0053] また、移動端末3は、上記のPMIPクライアント(A)13、PMIPクライアント(B)14、GW15がAAあるいはRAに格納した端末動作制御情報及びIP設定情報を読取る機能を備え、端末動作制御情報にしたがってPMIP動作とするか、CMIP動作とするかを決定する機能を備えている。さらに、移動端末3は、このとき読取ったIP設定情報によって自らのIP設定を行う機能を備えている。

[0054] アクセスネットワーク(A)200、アクセスネットワーク(B)300、アクセスネットワーク(C)400は、移動端末3がネットワーク100のノードと通信する際に経由するネットワークである。アクセスネットワークは、同一L2(Layer2)リンクでPMIPクライアント(A)13、PMIPクライアント(B)14、GW15に接続できる構成(すなわち、間にルータがな

い構成)であっても良いし、間にルータを配置した構成としても良い。

[0055] ここで、MIPv4のAS、AAに対し、それぞれ端末動作要求情報及び端末動作制御情報を格納する場合には、RFC3115で規定されたVSE(Vendor/Organization Specific Extension)を使用する例を示す。

[0056] VSEは、ベンダや組織固有の情報を格納する際に使用することができる。VSEは、さらに、CVSE(Critical Vendor/Organization Specific Extension)とNVSE(Normal Vendor/Organization Specific Extension)とが定義されている。

[0057] CVSEがRRQ、RRP、AA、ASといったMIPv4の信号に格納されていた場合には、これを認識しないMIPv4ノード[HA、FA(Foreign Agent)、CMIP端末、PMIPクライアント]が信号を破棄する。一方、NVSEの場合には、これを認識しないMIPv4ノードが、NVSE部位を無視して処理を継続する。

[0058] 本実施の形態では、本発明をサポートしないノードにおいてもMIPv4の処理が正しく行われるようにするため、NVSEを使用する。但し、CVSEを使用しても構わない。

[0059] NVSEのフォーマットを図2に示す。この拡張はMIPv4信号のいずれの拡張部位にも格納することができる。ここで、Typeは RFC3115によって「134」と定められている。Lengthには、Type及びLengthを含まない拡張データのバイト数が設定される。Vendor/Org－IDにはベンダあるいは組織を示す識別子が設定される。Vendor－NVSE－Typeには、ベンダあるいは組織毎に定義した拡張情報の種別を示す値が設定される。Vendor－NVSE－Valueには、Vendor－NVSE－Typeで示された拡張情報の値が設定される。

[0060] 図3は図1の移動端末3の構成例を示すブロック図である。図3において、移動端末3はCMIP動作手段31と、Simple IP動作手段32と、接続・動作検出手段33と、ポリシ蓄積手段34と、端末動作要求送信手段35と、端末動作情報受信手段36と、動作決定手段37とから構成されている。

[0061] CMIP動作手段31はRFC3344に規定されたMIPv4、あるいはRFC3775に規定されたMIPv6に準じた端末機能によってモビリティを実現する手段を備えている。

Simple IP動作手段32は一般的なIPv4、あるいはIPv6端末として動作するための手段を備えている。

[0062] 接続・移動検出手段33はレイヤ2の機能等によってリンクアップや異なるリンクに接続したことを検出する機能を備えている。これは、一般的な無線インタフェースでも実現することができ、本発明の特徴は含まない。ポリシ蓄積手段34はCMIP動作を要求するか、あるいはPMIP動作を要求するかのポリシ情報を格納する手段である。このポリシ情報はユーザが設定しても良いし、予め端末に設定しておくことにしても良い。

[0063] 端末動作要求送信手段35は接続・移動検出手段33によって、ネットワークへの接続、あるいは移動を検出した際に、移動端末3の動作モードをポリシ蓄積手段34に蓄積されたポリシ情報から取得し、その動作モードの情報を格納したAS(MIPv4の場合)、あるいはRS(MIPv6の場合)を送信する手段を備えている。

[0064] 例えば、移動端末3がCMIP動作を要求する場合には、CMIP動作要求を示す情報(例えば、CMIP動作要求に対応させたコード)をAS、あるいはRSに格納し、PMIP動作する場合には、一般的なAS、RSを送るようにすることが考えられる。

[0065] 但し、CMIP動作を要求しない場合でも、端末がCMIP機能を備えていることを通知したい場合には、CMIP機能を備えていることを示すコードをAS、あるいはRSに格納するようにしても良い。本実施の形態では、端末動作要求送信にAS(MIPv4の場合)、RS(MIPv6)を用いているが、それ以外のプロトコルを用いることとしても構わない。

[0066] 端末動作情報受信手段36はPMIPクライアント(A)13またはPMIPクライアント(B)14、あるいはGW15から送信されたAA、あるいはRAを受信し、それぞれに格納されかつPMIPクライアントが決定した移動端末3の動作モードや、決定された動作モードでの動作に必要な情報(HoA、HAアドレス等)を取得する手段を備えている。

[0067] 動作モードは、直接AA、RAに格納せずに前記決定された動作モードに必要な情報等から間接的に判定するようにしてもよい。本実施の形態では、端末動作情報受信にAA(MIPv4の場合)、RA(MIPv6)を用いているが、それ以外のプロトコルを用いることとしても構わない。

[0068] 動作決定手段37はAAあるいはRAで取得されたPMIPクライアントが決定した動作モードがCMIP動作であれば、CMIP動作手段31を有効化する機能を備えている。一方、動作決定手段37はPMIP動作であった場合、CMIP動作手段31を無効化し、Simple IP端末として動作するSimple IP動作手段32を有効化する機能を備えている。いずれの場合も、動作決定手段37はAAあるいはRAで取得した各動作に必要な情報の設定を行う機能も備える。

[0069] 尚、図示していないが、移動端末3は認証処理手段を有している。この認証処理手段はネットワークに接続したことが検出された際に、PMIPクライアントと認証処理を行う手段を備えている。この時、移動端末3は、RFC4282に規定されたNAI(Network Access Identifier)等の識別子を認証要求の中に格納する。

[0070] 図4は図1のPMIPクライアント(A)13の構成例を示すブロック図である。図4において、PMIPクライアント(A)13はPMIP動作手段131と、FA動作手段132と、端末動作要求受信手段133と、ポリシ通信手段134と、動作決定手段135と、認証処理手段136と、端末動作情報送信手段137とから構成されている。尚、図示していないが、PMIPクライアント(b)14はこのPMIPクライアント(A)13と同様の構成及び動作となっている。

[0071] PMIP動作手段131は上記の非特許文献2～4に規定された一般的なPMIPクライアントの動作をする手段を備えている。但し、PMIPクライアント(A)13が移動端末3に代わってRRQ(MIPv4の場合)、あるいはBU(MIPv6の場合)をHA11に送信する契機は、動作決定手段135から送信される通知を契機とする。

[0072] FA動作手段132はMIPv4の場合、RFC3344に規定されたFAとして動作するための機能も備えている。FA機能は、移動端末3にCMIP動作を許可した場合で、かつFAモードで動作させる際に使用される。但し、PMIP動作手段131が、移動端末3がFAモードで動作する際のように、RRQをFA経由でHA11に送る構成としても構わない。MIPv6の場合には、本手段は必要ない。

[0073] 端末動作要求受信手段133は移動端末3から送信されるAS(MIPv4の場合)あるいはRS(MIPv6)の場合を受信し、AS、あるいはRSに格納された移動端末3の動作要求情報を取得し、動作決定手段135に送信する機能を備えている。本実施の形

態では、端末が端末動作要求を送る手段としてAS(MIPv4の場合)、RS(MIPv6)を用いているが、それ以外のプロトコルを用いることとしても構わない。

[0074] ポリシ通信手段134はポリシ蓄積サーバ12に対し、接続した移動端末3の識別子をポリシ要求に格納し、そのポリシ要求を送信し、その結果として、ポリシ蓄積サーバ12から移動端末3をCMIP動作させるか、PMIP動作とするかのポリシ情報を取得する機能を備えている。ポリシ蓄積サーバ12に送信する情報には、当該PMIPクライアント(A)13の識別子や移動端末3の識別子以外にも、移動端末3の能力(CMIP端末を備えているか否か)等、他の情報を含むことにしても構わない。

[0075] PMIPクライアント識別子を含むことにした場合、ポリシ蓄積サーバ12は、どの移動端末かだけでなく、どのPMIPクライアント、あるいはどのアクセスネットワークに移動端末がいるかに依存してポリシを返すことができる。また、PMIPの識別子はPMIPの認証に用いることもできる。移動端末やPMIPの識別子としては、NAIを用いることができる。

[0076] 動作決定手段135はポリシ通信手段134で取得したポリシ情報と、端末動作要求[AS(RS)]受信手段133から取得した端末の動作要求情報とを基に移動端末3をPMIP動作とするか、CMIP動作とするかの動作モードを決定する機能を備えている。具体的な動作決定手順については後述する(図6参照)。

[0077] 移動端末3にCMIP動作させる場合で、かつMIPv4の場合には、さらに移動端末3がFAモードとCCoAモードとのどちらで動作すべきかを決定する。この決定のポリシとしては、例えば、PMIPクライアントがFAとしての機能を備えていればFAモード、FAとしての機能を備えていなければCCoAモードに決定する等とすることができる。但し、FAとしての機能を備えている場合でもCCoAモードを選択しても構わない。この場合、ポリシ蓄積サーバ12からFAモード、CCoAモードのどちらにするかのポリシを取得することで、どちらの動作にするかを決定するといった構成が考えられる。

[0078] 認証処理手段136は移動端末3から送られた識別子を用いて認証サーバ16と通信を行い、認証サーバ16から端末のHoA(CMIP端末の場合、Simple IP端末の場合にはHoAが単なるIPアドレス相当となる)、HA11のアドレス等の情報を取得する機能を備えている。この認証処理手段136は上記の非特許文献1にAuthenticat

orとして記載されている機能と同等ものであり、認証サーバ16としてRadiusサーバやDiameterサーバを使った構成が広く用いられている。

[0079] 端末動作情報送信手段137は動作決定手段135で決定した移動端末3の動作モードと、決定された動作モードで必要となるアドレス等の必要情報とをAA、RAに格納して移動端末3に送信する手段を備えている。アドレス等の必要情報は、認証処理手段16で取得した情報を用いる。本実施の形態では、PMIPクライアントが端末動作情報を送る手段としてAA(MIPv4の場合)、RA(MIPv6)を用いているが、それ以外のプロトコルを用いることとしても構わない。

[0080] 図5は図1のGW15の構成例を示すブロック図である。図5において、GW15はGW動作手段151と、FA動作手段152と、端末動作要求受信手段153と、認証処理手段154と、端末動作情報送信手段155とから構成されている。

[0081] GW動作手段151は一般的なルータとして動作する機能を備えている。また、GW動作手段151は移動端末3との間でIPsec(RFC4301に規定)によりセキュアな通信を確立する手段を備えていてもよい。いずれにせよ、GW動作手段151には本発明の特徴となる要素が含まれていない。FA動作手段152はMIPv4の場合、RFC3344に規定されたFAとして動作するための機能も備えている。FA機能は、移動端末3にFAモードで動作させる際に使用される。MIPv6の場合、このFA動作手段152は必要ない。

[0082] 端末動作要求受信手段153は移動端末3から送信されるAS(MIPv4の場合)あるいはRS(MIPv6の場合)を受信し、RS、あるいはASに格納された移動端末3の動作要求情報を取得する。但し、PMIP動作手段を備えていないため、"CMIP動作"が自動的に決定される。本実施の形態では、端末が端末動作要求を送る手段としてAS(MIPv4の場合)、RS(MIPv6)を用いているが、それ以外のプロトコルを用いることとしても構わない。

[0083] 認証処理手段154は移動端末3から送られた識別子を用いて認証サーバ16と通信を行い、認証サーバ16から端末のHoA(CMIP端末の場合、Simple IP端末の場合にはHoAが単なるIPアドレス相当となる)、HAのアドレス等の情報を取得する機能を備えている。この認証処理手段154は、上記の非特許文献1にAuthenticat

orとして記載されている機能と同等ものであり、認証サーバ16としてRadiusサーバやDiameterサーバを使った構成が広く用いられている。

[0084] 端末動作情報送信手段155は移動端末3にCMIP動作させるための動作モード"CMIP"と、認証処理手段154で取得した移動端末3に関連する情報(HoAやHAアドレス等の情報)、つまり移動端末3がCMIP端末として動作するために必要な情報を格納したAA、あるいはRAを移動端末3に送信する機能を備えている。移動端末3がMIPv4の場合には、さらに移動端末3がFAモードとCCoAモードとのどちらで動作すべきかも決定する。この決定のポリシとしては、例えば、GW15がFAとしての機能を備えていればFAモード、FAとしての機能を備えていなければCCoAモードに決定する等とすることができる。

[0085] 一方、GW15にFA機能を備えている場合にも、ポリシ蓄積サーバ12と通信する手段を備えることで、ポリシ蓄積サーバ12からFAモード、CCoAモードのどちらにするかのポリシを取得し、どちらの動作にするかを決定することにしても構わない。本実施の形態では、GW15が端末動作情報を送る手段としてAA(MIPv4の場合)、RA(MIPv6の場合)を用いているが、それ以外のプロトコルを用いることとしても構わない。

[0086] 図6は本発明の第1の実施の形態におけるNVSEが格納されたMIPv4信号を受信したMIPv4ノードの処理を示すフローチャートである。この図6を参照して、NVSEが格納されたMIPv4信号を受信したMIPv4ノードの処理について説明する。

[0087] MIPv4ノードは、MIPv4信号を受信した場合(図6ステップS1)、自ノードがNVSE Typeをサポートしているかどうかを判定し(図6ステップS2)、サポートしていなかった場合、NVSEを無視し、当該MIPv4信号にNVSEがなかったものとして処理を継続する(図6ステップS3)。

[0088] 一方、MIPv4ノードは、NVSE typeをサポートしている場合、次にVendor/Org-ID及びVendor-NVSE-Typeを認識可能であるかどうかを判定し(図6ステップS4)、いずかが認識できなかった場合、当該NVSEを無視し、当該NVSEがなかったものとして処理を継続する(図6ステップS5)。また、MIPv4ノードは、上記のステップS4の判定でいずれも認識可能であった場合、当該NVSEを適切に処理する(図

6ステップS6)。

[0089] このように、NVSEを使用することで、当該拡張を認識しないノードは、これを無視して当該拡張が格納された信号の処理を継続することが可能である。MIPv4においては、上述した端末動作要求情報、前記端末動作制御情報、IP設定情報といった情報を、NVSEを利用して送るものとする。

[0090] 一方、MIPv6の場合には、RAに移動端末の動作制御情報及びIP設定情報を格納するが、RAのoptionとして拡張情報を格納することで、拡張情報を認識するノードに情報を処理させ、それ以外のノードには当該情報を無視させることができる。

[0091] すなわち、この場合には、上記のMIPv4のNVSEを使用した場合と同様のことができる。したがって、MIPv6においては、上記の端末動作要求情報、前記端末動作制御情報、IP設定情報といった情報をRAのoptionを利用して送るものとする。

[0092] ここでは、移動端末3がMIPv4をサポートしており、HA11がMIPv4のHAであり、PMIPクライアント(A)13、PMIPクライアント(B)14がPMIPv4である場合を考慮し、また、図1の他のノードもIPv4に対応している状況を考慮する。

[0093] 本実施の形態において、ネットワーク100とアクセスネットワーク(A)200との境界に配置されたPMIPクライアント(A)13は、アクセスネットワーク(A)200の認証クライアント機能も備えている場合を想定する。さらに、PMIPクライアント(A)13は、FA機能も備えていてもよい、またFA機能を備えていない場合でも、FAのAS、AAをそれぞれ受信、送信する機能を備えているものとする。

[0094] 図7は本発明の第1の実施の形態において移動端末3がネットワークのポリシによってPMIP動作とする場合の動作を示すシーケンスチャートであり、図8は本発明の第1の実施の形態において移動端末3からの要求とネットワークのポリシとによって移動端末3の動作を決定する手順を示すフローチャートである。

[0095] また、図9は本発明の第1の実施の形態において移動端末3がアクセスネットワーク(A)200に接続した時にネットワークのポリシによってCMIP動作が許可される場合の動作を示すシーケンスチャートであり、図10は本発明の第1の実施の形態において移動端末3がアクセスネットワーク(C)400に接続した場合の動作を示すシーケンスチャートである。

[0096] はじめに、上記の図1及び図7を参照して、MIPv4機能を備えた移動端末3がアクセスネットワーク(A)200に接続した場合で、ネットワークのポリシによってPMIP動作とする場合の動作について説明する。

[0097] まず、移動端末3がアクセスネットワーク(A)200に接続すると、PMIPクライアント(A)13との間でアクセス認証処理を開始する。その際、移動端末3は認証のための秘密の情報と合わせて、ユーザを識別子するためのユーザ識別子をPMIPクライアント(A)13に送信する。

[0098] PMIPアクライアント(A)13は、移動端末3から受け取った認証情報を認証サーバ16に問合せることによって、移動端末3がアクセスの許可されたユーザかどうかを判定する。移動端末3がアクセスの許可されたユーザであると判定された場合、PMIPクライアント(A)13は、認証サーバ16から移動端末3に割当てるべきHoA(移動端末3がPMIP動作する場合には、HoAが移動端末3にとって単なるアドレスとみなされる)、アンカーとすべきHAアドレスをダウンロードする(図7のa1)。この時、上記の情報以外にも、HAに位置登録を行うための鍵情報等の他の情報もダウンロードするようにしても構わない。

[0099] 次に、移動端末3は、上記のa1の処理において、アクセス認証に成功した場合、ASをRFC1256に定義されたSolicitation Address(224. 0. 0. 2)、Mobile－Agents multicast group address(224. 0. 0. 11)、あるいはAddress(255. 255. 255. 255)宛に送信する(図7のa2)。

[0100] 移動端末3がCMIPを備えており、さらにCMIPとして動作することを希望する場合には、ASにCMIP要求を意味する情報を格納することができる。CMIP要求を意味する情報を送ることなく、CMIP要求の意思をPMIPクライアントに通知したい場合には、Address(224. 0. 0. 2)宛にASを送信した時にPMIP動作希望、Address(234. 0. 0. 11)宛にASを送信した時にCMIP動作希望、といったように宛先アドレスにPMIP要求、CMIP要求の対応関係を定義する方法とすることもできる。

[0101] 尚、移動端末3がCMIP動作を希望しない場合には、CMIP動作を希望しない情報をASに格納しても良いし、単なるASとしても良い。CMIP動作を希望しないもののネットワークに対し、CMIP機能が有ることを通知する場合には前者の方法とする。

[0102] 次に、PMIPクライアント(A)13は、ポリシ蓄積サーバ12に対してポリシ要求を送信する(図7のa3)。ここで、ポリシ蓄積サーバ12にPMIPクライアントやユーザ単位でのポリシ情報が蓄積されている場合には、PMIPクライアントの識別子と、上記のa1の処理で取得したユーザ識別子を送り、PMIPクライアントやユーザ個別のポリシ情報を要求することができる。尚、このa3の処理は、必ずしも上記のa2の処理の後である必要はない。

[0103] ポリシ蓄積サーバ12がa3の処理にてポリシ要求を受信した場合、PMIPクライアント(A)13に対し、予め設定されたポリシを通知する。ポリシ要求にPMIPクライアント識別子やユーザ識別子が格納されている場合には、それらに対応するポリシを検索し、検索された結果をPMIPクライアント(A)13に返す(図7のa4)。この動作例では、"PMIP強制"とのポリシをPMIPクライアント(A)13に応答する。

[0104] PMIPクライアント(A)13は、上記のa2の処理で得られた移動端末3の要求動作情報と、上記のa4の処理で得られたネットワークのポリシとによって、移動端末3の動作を決定する(図7のa5)。移動端末3の動作の決定手順については後述するが、ここでは、a4の処理において"PMIP強制"との応答を受信しているため、移動端末3に要求する動作はPMIPに決定される。

[0105] PMIPクライアント(A)13は、上記のa1の処理で取得した移動端末3のHoAと、自らのアドレスをCoAとして、同じくa1の処理で取得したHAアドレス宛にRRQを送信する(図7のa6)。

[0106] HA11は、a6の処理においてPMIPクライアント(A)13から送信されたRRQによって、HoA宛のパケットをカプセル化によってPMIPクライアント(A)13のアドレス宛に転送するためのBinding Cacheが作成される。その後、HA11は、RRQの応答としてRRPをPMIPクライアント(A)13に送信する(図7のa7)。

[0107] PMIPクライアント(A)13は、RRPを受信した場合、HA11からカプセル化によって送信されたパケットをデカプセルする用意をすると共に、移動端末3から送られたパケット(すなわち、送信元がHoAのパケット)をHA11宛にカプセル化して転送する用意を行う。

[0108] PMIPクライアント(A)13は、a7の処理でRRPを受信した後、AAを移動端末3に

送信する(図7のa8)。このAAには、NVSEを使って、PMIPとして動作すべきことを示す情報"PMIP"と、その際に移動端末3のアドレスとして使用すべきHoA、デフォルトGWアドレスを格納する。

[0109] さらに、AAの'H'フラグを'1'とし、さらにCoAフィールドにHAアドレスを格納することによって、本発明をサポートしないCMIP端末が当該AAを受信した場合にも、PMIP動作にすることができる。但し、この時点ではHAもHoAも未設定であり、Home戻りが検出できないため、RFC4433に規定された動的HA割当て手順が処理される。これによって、HoAを取得した時点でHome戻りが検出され、その後、PMIP動作(Simple IP端末としての動作)が可能となる。

[0110] このように、本発明をサポートしない移動端末の動作もPMIP動作となるように制御ができるが、本発明をサポート移動端末3と比較して信号の数や時間的に不利である。

[0111] 移動端末3は、a8の処理においてAAを受信した場合、AAに格納された"PMIP"情報によってPMIP動作することを決定する。さらに、移動端末3は、AAに格納されたHoA、デフォルトGWアドレス情報によってIP設定を完了する(図7のa9)。このa9の処理によって、アドレス設定が完了した後、移動端末3とCN2との間での通信が可能となる(図7のa10)。

[0112] ここで、a3の処理において、移動端末3からの要求とネットワークのポリシとによって移動端末3の動作を決定する手順について図8を参照して説明する。

[0113] まず、PMIPクライアント(A)13は自ノードのPMIP機能が有効かどうかを判定し(図8ステップS11)、PMIP機能が有効に動作していれば、図7のa4の処理で取得したネットワークポリシが"PMIP強制"であるかどうかを判定する(図8ステップS12)。ここで、ネットワークポリシが"PMIP強制"であった場合、PMIPクライアント(A)13は移動端末3の動作を"PMIP"とすることに決定する(図8ステップS13)。

[0114] ネットワークポリシが"PMIP強制"でなかった場合、PMIPクライアント(A)13はネットワークポリシが"CMIP許可"かどうかを判定する(図8ステップS14)。PMIPクライアント(A)13はネットワークポリシが"CMIP許可"であった場合、図7のa1の処理で取得した移動端末3の動作要求が"CMIP動作要求"かどうかを判定する(図8ステップ

S15)。

[0115] “PMIP動作要求”、あるいは、例えば、図7のa1の処理で単なるASを受信した場合等のように、CMIP要求動作が明示的に示されていない場合、PMIPクライアント(A)13は移動端末3の動作を“PMIP”とすることに決定する(図8ステップS13)。

[0116] “CMIP動作要求”が示されていた場合、PMIPクライアント(A)13は移動端末3の動作を“CMIP”に決定する(図8ステップS16)。上記のステップS14の処理で、ネットワークポリシが“CMIP許可”でなかった場合も、PMIPクライアント(A)13は移動端末3の動作を“CMIP”に決定する(図8ステップS16)。

[0117] この動作例において、PMIPクライアント(A)13では、上記のステップS11の判定結果が基本的に“YES”(PMIP機能有効)となる。しかしながら、PMIPクライアント(A)13において何らかの理由でPMIP機能が動作していない場合や、GW15のようにPMIP機能を備えていない場合には、上記のステップS11の判定結果が“NO”(PMIP機能無効)となるので、移動端末3の動作が“CMIP”に決定される。

[0118] 次に、MIPv4機能を備えた移動端末3がアクセスネットワーク(A)200に接続した時にネットワークのポリシによってCMIP動作が許可される場合の動作について図9を参照して説明する。

[0119] 図9のb1の処理からb4の処理までは、図7に示したPMIP強制の場合のa1の処理からa4の処理と同じである。但し、b4の処理でポリシ蓄積サーバ12から応答されるポリシは“CMIP許可”である。

[0120] 次に、PMIPクライアント(A)13は、図8で説明した移動端末3の動作の決定手順によって移動端末3の動作を“CMIP動作”に決定する(図9のb5)。PMIPクライアント(A)13は、このb5の処理で移動端末3の動作を“CMIP動作”に決定した場合、“PMIP強制”の場合とは異なり、RRQをHAに送ることはせず、AAを移動端末3に送信する(図9のb6)。

[0121] このAAには、NVSEを使って、移動端末3がCMIPとして動作すべきことを示す情報“CMIP”と、CMIPとして動作するために必要なHoA、CoA、及びFAモードとCCoA(Colocated－CoA)とのどちらで動作すべきかの情報を格納する。このFAモードとCCoAとのどちらで動作すべきかの情報は、ポリシ蓄積サーバ12から取得する構

成としても構わない。

[0122] さらに、AAの‘H’フラグを‘1’とし、さらにCoAフィールドにHAアドレスを格納する。この設定によって、本発明をサポートしないCMIP端末が当該AAを受信した場合には、PMIP動作にすることができる。但し、上記のように、この場合にはPMIP動作によって通信可能となるまでに、より多くの信号と時間とが使われる。

[0123] 移動端末3は、b6の処理においてAAを受信した場合、AAに格納された“CMIP”情報によってCMIP動作とすることに決定する。また、移動端末3は、AAに格納されたHoA, CoAによってIP設定を行い、さらにFA／CCoA情報によってFAモードで動作するか、CCoAモードで動作するかを決定する(図9のb7)。尚、図9に示す例ではFAモードで動作する場合を考慮する。

[0124] 移動端末3は、b7の処理でCMIP・FAモードで動作することを決定した後、HoA、CoAをHA11に登録するために、PMIPクライアント(A)13にRRQを送信する。PMIPクライアント(A)13は、移動端末3からのRRQをHA11に転送する〔この場合、PMIPクライアント(A)13がFA相当の動作を行う〕(図9のb8)。

[0125] HA11は、b8の処理においてRRQを受信した場合、HoA宛のパケットをカプセル化によってPMIPクライアント(A)13のアドレス宛に転送するためのBinding Cacheを作成する。その後、HA11は、RRQの応答としてRRPをPMIPクライアント(A)13に送信する。

[0126] PMIPクライアント(A)13は、HA11からのRRPを受信した場合、HA11からカプセル化によって送信されたパケットをデカプセルする用意を行うと共に、移動端末3から送られたパケット(すなわち、送信元がHoAのパケット)をHA11宛にカプセル化して転送する用意を行い、最後にRRPを移動端末3に転送する(図9のb9)。このb9の処理においてHA11への位置登録が完了した後、移動端末3とCN2との間での通信が可能となる(図9のb10)。

[0127] 次に、図1と図10とを参照して、移動端末3がアクセスネットワーク(C)400に接続した場合の動作について説明する。図10において、c1, c2の処理は、図7に示すPMIP強制の場合のa1, a2の処理と同じである。

[0128] GW15は、PMIPクライアント機能を備えていないため、c2の処理でASを受信する

と, CMIP動作が要求されているか否かに関わらず、移動端末3の動作を"CMIP"に決定する(図10のc3)。

[0129] 次に、GW15は, AAを移動端末3に送信する(図10のc4)。このAAには、NVSEを使って移動端末3がCMIPとして動作すべきことを示す情報"CMIP"と、CMIPとして動作するために必要なHoA, CoA、及びFAモードとCCoA(Colocated－CoA)とのどちらで動作すべきかの情報を格納する。

[0130] 例えば、GW15がFA機能を備えている場合にはFAモードとし、FA機能を備えていない場合にはCCoAとすることができる。ここでは、CCoAを選択した場合を考慮する。さらに、FAモードを選択した場合には、AAの'F'フラグを'1'とし、さらにCoAフィールドにGWアドレス(すなわち, FAアドレス)を格納する。

[0131] 一方, CCoAモードを選択した場合には, AAの'F'フラグと'R'フラグとを'1'とし、さらにCoAフィールドにはHA11からGW配下のネットワークに割り当てられた到達性のあるアドレスを格納する。この設定によって、本発明をサポートしないCMIP端末が当該AAを受信した場合でも、それぞれFAモードでの動作, CCoAモードでの動作とすることができる。いずれの場合も、その後、HoAを取得する処理が必要となる。CCoAモードで動作する場合には、本発明をサポートしないCMIP端末は、さらにHAアドレスを事前に設定しておくか、なんらかの方法でHAのアドレスを取得しなければならない。

[0132] 移動端末3は, c5の処理においてAAを受信した場合、AAに格納された"CMIP"情報及びCCoAモードで動作することを示す情報によってCCoAモードでCMIP動作することを決定する。また、移動端末3は, AAに格納されたHoA、CoAによって、IP設定を行い、同じくAAに格納されたHAアドレスを取得し, これを保持しておく。

[0133] 移動端末3は、上記のc5の処理で設定したHoAとCoAとを対応付けるためのRRQを、同じくc5の処理で取得したHAアドレスに対して送信する(図10のc6)。

[0134] HA11は、c6の処理においてRRQを受信した場合、HoA宛のパケットをカプセル化によって移動端末3のCoAに転送するためのBinding Cacheを作成する。その後、HA11は、RRQの応答としてRRPを移動端末3に送信する。移動端末3は、RRPを受信した場合, HA11からカプセル化によって送信されたパケットをデカプセル

する用意を行う(図10のc7)。また、移動端末3が送信するパケットをカプセル化してHA11に送るための用意を行う。このc7の処理においてHA11への位置登録が完了した後、移動端末3とCN2との間での通信が可能となる(図10のc8)。

[0135] 次に、移動端末3がハンドオーバする状況を考慮する。まず、移動端末3が、アクセスネットワーク(A)200からアクセスネットワーク(B)300へハンドオーバする場合のように、PMIPをサポートするアクセスネットワークへハンドオーバするケースについて考える。

[0136] この場合の動作は、図7あるいは図9のアクセスネットワーク(A)200に接続した場合と同じとして構わない。この場合、図7、図9のPMIPクライアントは、PMIPクライアント(B)14となる。但し、移動端末3がアクセスネットワーク(B)300に接続した際にPMIPクライアントや移動端末3がRRQを送信するために必要な情報(HoA、HAアドレス、HAへの位置登録をセキュアに行うための鍵情報等)を認証サーバ16に問合せるのではなく、ハンドオーバ元のPMIPクライアント(A)13から取得する方式としても構わない。

[0137] この時, ハンドオーバ元のPMIPクライアントがどれであるかは、アクセスネットワークに配置されたBS(Base Station)から通知することができる。あるいは、上記の非特許文献1に示されたASN Functional Entityのようなものを利用することができる。しかしながら、具体的な方法は, 本発明に直接関係がないので、その詳細な説明については省略する。

[0138] また、ハンドオーバの場合には、移動端末3はHoAを既に取得しており, このHoAは通信を継続する間、保持し続けるアドレスであるため, a8の処理あるいはb6の処理においてAAに含める情報の中からHoAを省略しても構わない。

[0139] 次に、移動端末3が、アクセスネットワーク(B)300からアクセスネットワーク(C)400へハンドオーバする場合のように、PMIPがサポートしないアクセスネットワークへハンドオーバするケースについて考える。

[0140] この場合の動作は、図10のアクセスネットワーク(C)400に接続した場合と同じとして構わない。但し、移動端末3がアクセスネットワーク(C)400に接続した際に, 移動端末3がRRQを送信するために必要な情報(HoA、HAアドレス、HAへの位置登録

をセキュアに行うための鍵情報等)を認証サーバ16に問合せるのではなく、ハンドオーバ元のPMIPクライアント(B)14から取得する方式としても構わない。また、ハンドオーバの場合には、移動端末3はHoAを既に取得しており、このHoAは通信を継続する間, 保持し続けるアドレスであるため、e4の処理においてAAに含める情報の中からHoAを省略しても構わない。

[0141] 本実施の形態では、移動端末3とPMIPクライアント(A)13、PMIPクライアント(B)14、GW15との間での情報のやり取りにAS、AAを利用することとしたが、必ずしもAS、AAである必要はなく, 例えば, IPCP(Internet Protocol Control Protocol configure－request, IPCP configure－Ackを使用することができる。

[0142] 本実施の形態で示した拡張情報を含むASの送信や拡張情報の含まれたAAの解釈をサポートする移動端末3は, CMIP機能を備えていることを想定しているが、これは、現実的にはSimple IP端末が一般的なIP端末と異なる動作をするとの前提を設けづらいためである。しかしながら、Simple IP端末が本発明の移動端末3であっても構わない。

[0143] また、本実施の形態では、移動端末3がCMIPで動作する要求を送信する場合を考慮したが、例えば、PMIP機能を提供可能なアクセスネットワークに接続した場合には必ずPMIPで動作するとのポリシとすれば, 移動端末3の送信するASは、本発明の拡張要素を含まない、一般的なASとしても構わない。この場合には、ポリシ蓄積サーバ12への問合せも省略することができる。

[0144] さらに, 本実施の形態ではAAによって, 移動端末3に通知するIP情報としてDNSサーバアドレスを含めていないが、DNSサーバアドレスを含めるようにしても構わない。

[0145] このように, 本実施の形態では、PMIPクライアントやGW15が、CMIP機能を備えた移動端末3の動作を制御するための情報を移動端末に通知することで、ネットワークの管理者のポリシにしたがって移動端末3がPMIP動作するか、CMIP動作するかを高速に決定することができる。

[0146] また、本実施の形態では、PMIPクライアントやGW15が、移動端末3の動作を制御するための情報とともに、それぞれの動作で必要となるIP設定情報も同時に移動

端末3に通知することで、移動端末3が動作の決定と同時に、IP設定をも完了することができ、高速に通信可能な状態にすることができる。

[0147] さらに、本実施の形態では、移動端末3もPMIP動作、CMIP動作のいずれかを希望する情報をネットワーク100に通知し、ネットワーク100が移動端末3の動作の希望と、ネットワーク100側のポリシの双方を鑑みて移動端末3の動作を制御することが可能となる。

[0148] 移動端末3の希望によってCMIP動作とした場合には、移動端末3側の意図によって、サービス毎に異なるHAをアンカーとして選択したり、移動端末3が複数の通信I／F(Interface)を備えていた場合に移動端末3がアクセスネットワークに接続した契機ではなく、端末を使用するユーザの意図するタイミング等で異なるI／Fへハンドオーバさせることが可能となる。異なるI／Fを同時に使用することも考えられる。

[0149] さらに、本実施の形態では、移動端末3の動作を制御するための情報や、MIP動作、CMIP動作のいずれかを希望する情報を一般的に使用されるプロトコルの信号(MIPv4の場合には、AS、AA)に格納することで、一般的な信号と本発明の情報を送るための信号とを別途送る必要がなく、通信帯域の効率利用を実現することができる。

[0150] さらにまた、本実施の形態では、一般的な信号の(本発明による拡張情報部位ではない)基本情報部分に本発明をサポートしていない一般的なCMIP端末の動作も制御可能な情報を格納するとで、本発明をサポートしないCMIP端末の動作をも制御することができる。

[0151] したがって、本発明のPMIPクライアントは、本発明をサポートするCMIP端末、本発明をサポートしないCMIP端末、Simple IP端末を効率的に収容することが可能であり、現存するシステム、あるいは現在検討が進められているシステムへの適用が容易である。

[0152] 図11は本発明の第2の実施の形態による移動端末管理システムの動作を示すシーケンスチャートであり、図12は本発明の第2の実施の形態において移動端末3がアクセスネットワーク(A)200に接続した場合にネットワークのポリシによってCMIP動作が許可された場合の動作を示すシーケンスチャートであり、図13は本発明の第2

の実施の形態において移動端末3がアクセスネットワーク(C)400に接続した場合の動作を示すシーケンスチャートである。

[0153] 本発明の第2の実施の形態による移動端末管理システムの構成は、図1に示す本発明の第1の実施の形態による移動端末管理システムと同様の構成となっている。但し、本発明の第2の実施の形態では、移動端末3がMIPv6をサポートしており、HA11がMIPv6のHAであり, PMIPクライアント(A)13, PMIPクライアント(B)14がPMIPv6である場合を考慮している。また、図1の他のノードもIPv6に対応している状況を考慮している。

[0154] 本実施の形態において, ネットワーク100とアクセスネットワーク(A)200との境界に配置されたPMIPクライアント(A)13は、アクセスネットワーク(A)200の認証クライアント機能も備えている場合を想定する。最初に, 移動端末3がアクセスネットワーク(A)200に接続した場合で, ネットワークのポリシによってPMIP動作とする場合の動作を、図1と図11とを参照して説明する。

[0155] 図11には、移動端末3にPMIP動作を強制する場合の動作例を示しているが、d1の処理からd10の処理については、図7のa1の処理からa10の処理に対応しており、ほとんど同じ動作となる。但し、d2, d8の処理においてAS、AAではなく、RS、RAを使用している点が異なる。また、RAに格納する情報も異なる。

[0156] 尚、d6, d7の処理におけるBU(Binding Update)、BA(Binding Acknowledge)も図7のRRQ、RRPとは異なるが、HA11に対し、HoAとCoAとを対応付けた情報の保持とHoA宛のパケットをカプセル化によってCoA宛に転送することを要求する点で、それらの機能は同じである。

[0157] MIPv6ではAS, AAは使用せず、移動検出やCoA決定のためにRS、RAを使用する。したがって、本実施の形態では、RS, RAのoption部分に情報を格納する。このRS、RAのoption部に情報を格納した場合、これを認識できないノードは当該optionを無視する動きになり、MIPv4の場合のNVSEと同様の使い方が可能である。

[0158] 図11のd8の処理において、RAには、option部分にPMIPでの動作を指定するための"PMIP"との情報を格納する。そして、RAの基本情報部分には移動端末3がHoAを生成するためのHome prefixとPMIPクライアントのRAを送信したL/FのLin

klayerアドレスを格納する。

[0159] ここで、基本部分にHome prefixの情報を格納することで、移動端末3が、本発明の端末制御情報に対応していない場合でもPMIP動作とすることができる。但し、この場合、移動端末3はRFC3775に記述されたDynamic Home Agent Address Discoveryや事前に設定しておく等の方法でHAを発見し、さらにHoAを取得する処理が必要である。その処理の後、HoAを取得できた時点で、ようやく移動端末3にHome戻りを検出することができ、ここではじめてPMIP動作が可能となる。

[0160] このように、ネットワークと交換する信号の数や時間は、本発明の端末制御情報に対応した移動端末よりも多くかかることになる。

[0161] 図11のd9の処理では、移動端末3が、d8の処理での"PMIP"通知によって、PMIP動作することに決定する。この時、同じくd8の処理で取得したHome prefixと自らのI/F識別子[例えば、MAC(Media Access Control)アドレス]からアドレスを生成し、自らのアドレスとする。

[0162] また、デフォルトGWのアドレスとしてRAの送信元アドレスを設定し、さらにPMIPクライアントのRA送信元のLinklayerアドレスも取得し、RAの送信元アドレス(PMIPクライアントのアドレス)と対応付けて記憶しておく。これによって、初めてデフォルトGWにパケットを送る際にND(Neighbor Discovery)プロトコルの処理によってLinklayerアドレスを取得するための信号交換を省くことができ、効率が良くなる。但し、Linklayerアドレスを RAに含めないようにしても構わない。一方で、RAにDNSサーバアドレスを含む構成としても構わない。

[0163] d9の処理で、移動端末3はPMIP動作することを決定し、さらに必要なIPアドレス設定も完了した後、CN2との通信を開始することができる(図11のd10)。

[0164] 次に、移動端末3がアクセスネットワーク(A)200に接続した場合で、ネットワークのポリシによってCMIP動作を許可された場合の動作を、図1と図12とを参照して説明する。

[0165] 図12のe1からe3までの処理は、図11のd1からd3までの処理と同じである。e4の処理では、図11のd4の処理とは違い、ポリシ蓄積サーバ12から"CMIP許可"が応答される。

[0166] e5の処理では、e2の処理での移動端末3から送られた"CMIP要求"と、e4の処理で応答された"CMIP許可"の情報と、図8に示したPMIP、CMIP動作判定手順にしたがって移動端末3の動作を決定する。結果としては、"CMIP動作"が選択される。

[0167] e6の処理ではRAを移動端末3に送信する。RAには、option部分にCMIPでの動作を指定するための"CMIP"情報と、移動端末3がCMIP動作をするために必要なHAアドレスとForeign prefixとを格納する。そして、RAの基本情報部分には移動端末3がHoAを生成するためのHome prefixとPMIPクライアントのRAを送信したI/FのLinklayerアドレスとを格納する。この場合、上述したように、本発明の端末制御情報を認識できない移動端末はPMIP動作となる。

[0168] e6の処理で受信したRAの"CMIP"情報によって、移動端末3はCMIP動作することを決定する(図12のe7)。さらに、移動端末3はRAに格納されたHome prefix、Foreign prefixからそれぞれHoA、CoAを生成し、自らにHoA、CoAを割当てた後、同じくRAに格納されたHAアドレス宛にHoAとCoAとの対応を登録するためにBUを送信する(図12のe8)。移動端末3はBUの応答としてBAを受信し(図12のe9)、その後、CN2との通信を開始することができる(図12のe10)。

[0169] 次に、移動端末3がアクセスネットワーク(C)400に接続した場合の動作を、図1と図13とを参照して説明する。図13のf1, f2の処理は、図11のd1, d2の処理と同じである。

[0170] GW15は、PMIPクライアント機能を備えていないため、f2の処理でRSを受信すると、CMIP動作が要求されているか否かに関わらず、移動端末3の動作をCMIPに決定し(図13のf3)、RAを移動端末3に送信する(図13のf4)。RAには、option部分にCMIPでの動作を指定するための"CMIP"情報と、移動端末3がCMIP動作をするために必要なHAアドレスと、移動端末3がHoAを生成するために用いるHome prefixとを格納する。そして、RAの基本情報部分には移動端末3がCoAを生成するためのForeign prefixとPMIPクライアントのRAを送信したI/FのLinklayerアドレスとを格納する。

[0171] この場合、本発明の端末制御情報を認識できない移動端末はCMIP動作となる。但し、この場合、移動端末3はRFC3775に記述されたDynamic Home Agent

Address Discoveryや事前に設定しておく等の方法でHAを発見し、さらにHoAを取得する処理が必要となる。

[0172] (4の処理で受信したRAの"CMIP"情報によって、移動端末3はCMIP動作することを決定する(図13のf5)。さらに、移動端末3はRAに格納されたHome prefix、Foreign prefixからそれぞれHoA、CoAを生成し、自らにHoA、CoAを割当てた後、同じくRAに格納されたHAアドレス宛にHoAとCoAとの対応を登録するためにBUを送信する(図13のf6)。移動端末3はBUの応答としてBAを受信し(図13のf7)、その後、CN2との通信を開始することができる(図13のf8)。

[0173] 次に、移動端末3がハンドオーバする状況を考慮する。まず、移動端末3が、アクセスネットワーク(A)200からアクセスネットワーク(B)300へハンドオーバする場合のように、PMIPをサポートするアクセスネットワークへハンドオーバするケースについて考える。

[0174] この場合の動作は、図11あるいは図12のアクセスネットワーク(A)200に接続した場合と同じとして構わない。この場合、図11、図12のPMIPクライアントは、PMIPクライアント(B)14となる。但し、移動端末3がアクセスネットワーク(B)300に接続した際にPMIPクライアントや移動端末3がBUを送信するために必要な情報(HoA、HAアドレス、HAへの位置登録をセキュアに行うための鍵情報等)を認証サーバ16に問合せるのではなく、ハンドオーバ元のPMIPクライアント(A)13から取得する方式としても構わない。

[0175] この時、ハンドオーバ元のPMIPクライアントがどれであるかは、アクセスネットワークに配置されたBS(Base Station)から通知することができる。あるいは、上記の非特許文献1に示されたASN Functional Entityのようなものを利用することができる。しかしながら、具体的な方法は、本発明に直接関係がないので、その詳細な説明については省略する。

[0176] また、ハンドオーバの場合には、移動端末3がHoAを既に取得しており、このHoAは通信を継続する間、保持し続けるアドレスであるため、図11のd8の処理においてRAに含める情報の中からHome prefixを省略しても構わない。

[0177] 次に、移動端末3が、アクセスネットワーク(B)300からアクセスネットワーク(C)400

へハンドオーバする場合のようにPMIPがサポートしないアクセスネットワークへハンドオーバするケースについて考える。

[0178] この場合の動作は、図13のアクセスネットワーク(C)6に接続した場合と同じとして構わない。但し、移動端末3がアクセスネットワーク(C)6に接続した際に、移動端末3がBUを送信するために必要な情報(Home prefix、HAアドレス、HAへのBUをセキュアに行うための鍵情報等)を認証サーバ16に問合せるのではなく、ハンドオーバ元のPMIPクライアント(B)14から取得する方式としても構わない。

[0179] また、ハンドオーバの場合には、移動端末3がHoAを既に取得しており、このHoAは通信を継続する間、保持し続けるアドレスであるため、図13のf4の処理においてRAに含める情報の中からHome prefixを省略しても構わない。

[0180] 本実施の形態では、移動端末3とPMIPクライアント(A)13、PMIPクライアント(B)14、GW15との間での情報のやり取りに、RS、RAを利用することとしているが、必ずしもRS、RAである必要はなく、例えば、IPCPv6 configure－request、IPCPv6 configure－Ackを使用することができる。

[0181] 本実施の形態で示した拡張情報を含むRSの送信や拡張情報が含まれたRAの解釈をサポートする移動端末3は、CMIP機能を備えていることを想定しているが、これは、現実的にはSimple IP端末が一般的なIP端末と異なる動作をするとの前提を設けづらいためである。しかしながら、Simple IP端末が本発明の移動端末3としての機能を備えていても構わない。

[0182] また、本実施の形態では、移動端末3がCMIPで動作する要求を送信する場合を考慮しているが、例えば、PMIP機能を提供可能なアクセスネットワークに接続した場合は必ずPMIPで動作するとのポリシとすれば、移動端末3が送信するRSは、本発明の拡張要素を含まない一般的なRSとしても構わない。この場合は、ポリシ蓄積サーバ12への問合せも省略することができる。

[0183] さらに、本実施の形態ではRAによって、移動端末3に通知するIP情報としてDNS(Domain Name System)サーバアドレスを含めていないが、DNSサーバアドレスを含めるようにしても構わない。

[0184] このように、本実施の形態では、移動端末3の動作要求情報や端末の動作制御情

報をそれぞれ、RS、RAに格納することで、MIPv6を備えた移動端末3においても高速なPMIP、CMIP動作の決定とそれぞれでのIP設定が可能となる。

[0185] 上述した本発明は、3GPP(3rd Generation Partnership Project)や3GPP2、あるいはWiMAX Forumで策定しているモバイル向けネットワークのように、PMIP端末とCMIP端末とが共存する環境において有効である。

[0186] また、標準化団体によって、PMIPクライアントやモビリティアンカーの名称やモビリティ機能以外の機能配置等、異なる点はあるが、モビリティ機能はMIPv4、MIPv6プロトコルを使用しており、本発明の適用が可能である。さらに、先の標準化団体が検討しているネットワークは、数千万台程度の端末を収容するネットワークとなる。このようなネットワークでは信号交換の効率化やネットワークへの接続時、あるいはハンドオーバ時に迅速なサービス提供を可能とすることが極めて重要な課題となり、本が詰め胃によると、この課題の解決ないしは低減が可能となる。

[0187] 本出願は、2007年2月13日に出願された日本出願特願2007−031508号を基礎とする優先権を主張し、その開示の全てをここに取り込む。

請求の範囲

[1] PMIP[Proxy MIP(Mobile Internet Protocol)]機能を持つネットワーク機器が配備されたネットワークにおいて移動端末の動作を制御する移動端末管理システムであって,

前記ネットワーク側から前記移動端末側に前記PMIP機能を提供可能か否かを通知することを特徴とする移動端末管理システム。

[2] 前記移動端末は、自端末がCMIP(Client MIP)機能を備えている時に前記CMIP機能の通知及び前記CMIP機能の使用の要求のうちの少なくともいずれかを前記ネットワーク側に送信することを特徴とする請求項1記載の移動端末管理システム。

[3] 前記ネットワーク機器は、前記移動端末からの要求、前記ネットワーク側のポリシ、前記PMIP機能の有無、それらの組合せのいずれかによって前記移動端末をPMIP動作とするかCMIP動作とするかを決定する手段と、その決定した動作を前記移動端末に通知する手段とを含むことを特徴とする請求項2記載の移動端末管理システム。

[4] 前記ネットワーク機器は、前記決定した動作を前記移動端末に通知するとともに、前記決定した動作を実行する上で必要となる設定情報を前記移動端末に送出し、

前記移動端末は、前記PMIP動作とするかCMIP動作とするかを決定した際に、その動作に必要な設定を行うことを特徴とする請求項3記載の移動端末管理システム。

[5] 前記ネットワーク側と前記移動端末との間の通知を、MIP端末及びIP(Internet Protocol)端末において使用するプロトコルの拡張部分にかつ当該拡張部分を解釈できないノードにおいてその拡張部分が無視される性質を有する拡張部分に格納することを特徴とする請求項1から請求項4のいずれか記載の移動端末管理システム。

[6] 前記ネットワーク側と前記移動端末との間の信号に、前記通知をやり取りしないCMIP端末の動作を制御するための情報を格納することを特徴とする請求項5記載の移動端末管理システム。

[7] 前記プロトコルは、MIPv4(Mobile Internet Protocol version 4)を使用する場合に、AS(Agent Solicitation)、AA(Agent Advertisement)であることを特徴とする請求項5または請求項6記載の移動端末管理システム。

[8] 前記プロトコルは、MIPv6(Mobile Internet Protocol version 7)を使用する場合に、RS(Router Solicitation)、RA(Router Advertisement)であることを特徴とする請求項5または請求項6記載の移動端末管理システム。

[9] 前記ネットワーク機器は、3GPP－A(3GPP－Anchor)、MME(Mobility Management Entity)／UPE(User Plane Entity)、PDG(Packet Data Gateway)、PDSN(Packet Data Serving Node)、PDIF(Packet Data Interworking Function)、ASN(Access Service Network)－GW(GateWay)のいずれかであることを特徴とする請求項1から請求項8のいずれか記載の移動端末管理システム。

[10] PMIP[Proxy MIP(Mobile Internet Protocol)]機能を持つネットワーク機器が配備されたネットワークにおいて移動端末の動作を制御する移動端末管理システムを構成するネットワーク機器であって、

前記移動端末側に対して前記PMIP機能を提供可能か否かを通知することを特徴とするネットワーク機器。

[11] 前記移動端末がCMIP(Client MIP)機能を備えている時に前記CMIP機能の通知及び前記CMIP機能の使用の要求のうちの少なくともいずれかを前記移動端末から受信することを特徴とする請求項10記載のネットワーク機器。

[12] 前記移動端末からの要求、前記ネットワーク側のポリシ、前記PMIP機能の有無、それらの組合せのいずれかによって前記移動端末をPMIP動作とするかCMIP動作とするかを決定する手段と、その決定した動作を前記移動端末に通知する手段とを含むことを特徴とする請求項11記載のネットワーク機器。

[13] 前記決定した動作を前記移動端末に通知するとともに、前記決定した動作を実行する上で必要となる設定情報を前記移動端末に送出し、

前記移動端末が、前記PMIP動作とするかCMIP動作とするかを決定した際に、その動作に必要な設定を行うことを特徴とする請求項12記載のネットワーク機器。

[14] 前記移動端末との間の通知を、MIP端末及びIP(Internet Protocol)端末において使用するプロトコルの拡張部分にかつ当該拡張部分を解釈できないノードにおいてその拡張部分が無視される性質を有する拡張部分に格納することを特徴とする

請求項10から請求項13のいずれか記載のネットワーク機器。

[15] 前記移動端末との間の信号に、前記通知をやり取りしないCMIP端末の動作を制御するための情報を格納することを特徴とする請求項14記載のネットワーク機器。

[16] 前記プロトコルは、MIPv4(Mobile Internet Protocol version 4)を使用する場合に、AS(Agent Solicitation)、AA(Agent Advertisement)であることを特徴とする請求項14または請求項15記載のネットワーク機器。

[17] 前記プロトコルは、MIPv6(Mobile Internet Protocol version 7)を使用する場合に、RS(Router Solicitation)、RA(Router Advertisement)であることを特徴とする請求項14または請求項15記載のネットワーク機器。

[18] 3GPP－A(3GPP－Anchor)、MME(Mobility Management Entity)／UPE(User Plane Entity)、PDG(Packet Data Gateway)、PDSN(Packet Data Serving Node)、PDIF(Packet Data Interworking Function)、ASN(Access Service Network)－GW(GateWay)のいずれかであることを特徴とする請求項10から請求項17のいずれか記載のネットワーク機器。

[19] PMIP[Proxy MIP(Mobile Internet Protocol)]機能を持つネットワーク機器が配備されたネットワークにおいて移動端末の動作を制御する移動端末管理システムに用いる移動端末動作制御方法であって、

前記ネットワーク側から前記移動端末側に前記PMIP機能を提供可能か否かを通知することを特徴とする移動端末動作制御方法。

[20] 前記移動端末がCMIP(Client MIP)機能を備えている時に前記CMIP機能の通知及び前記CMIP機能の使用の要求のうちの少なくともいずれかを前記移動端末から前記ネットワーク側に送信することを特徴とする請求項19記載の移動端末動作制御方法。

[21] 前記ネットワーク機器が、前記移動端末からの要求、前記ネットワーク側のポリシ、前記PMIP機能の有無、それらの組合せのいずれかによって前記移動端末をPMIP動作とするかCMIP動作とするかを決定する処理と、その決定した動作を前記移動端末に通知する処理とを実行することを特徴とする請求項20記載の移動端末動作制御方法。

[22] 前記ネットワーク機器が、前記決定した動作を前記移動端末に通知するとともに、前記決定した動作を実行する上で必要となる設定情報を前記移動端末に送出し、

前記移動端末が、前記PMIP動作とするかCMIP動作とするかを決定した際に、その動作に必要な設定を行うことを特徴とする請求項21記載の移動端末動作制御方法。

[23] 前記ネットワーク側と前記移動端末との間の通知を、MIP端末及びIP(Internet Protocol)端末において使用するプロトコルの拡張部分にかつ当該拡張部分を解釈できないノードにおいてその拡張部分が無視される性質を有する拡張部分に格納することを特徴とする請求項19から請求項22のいずれか記載の移動端末動作制御方法。

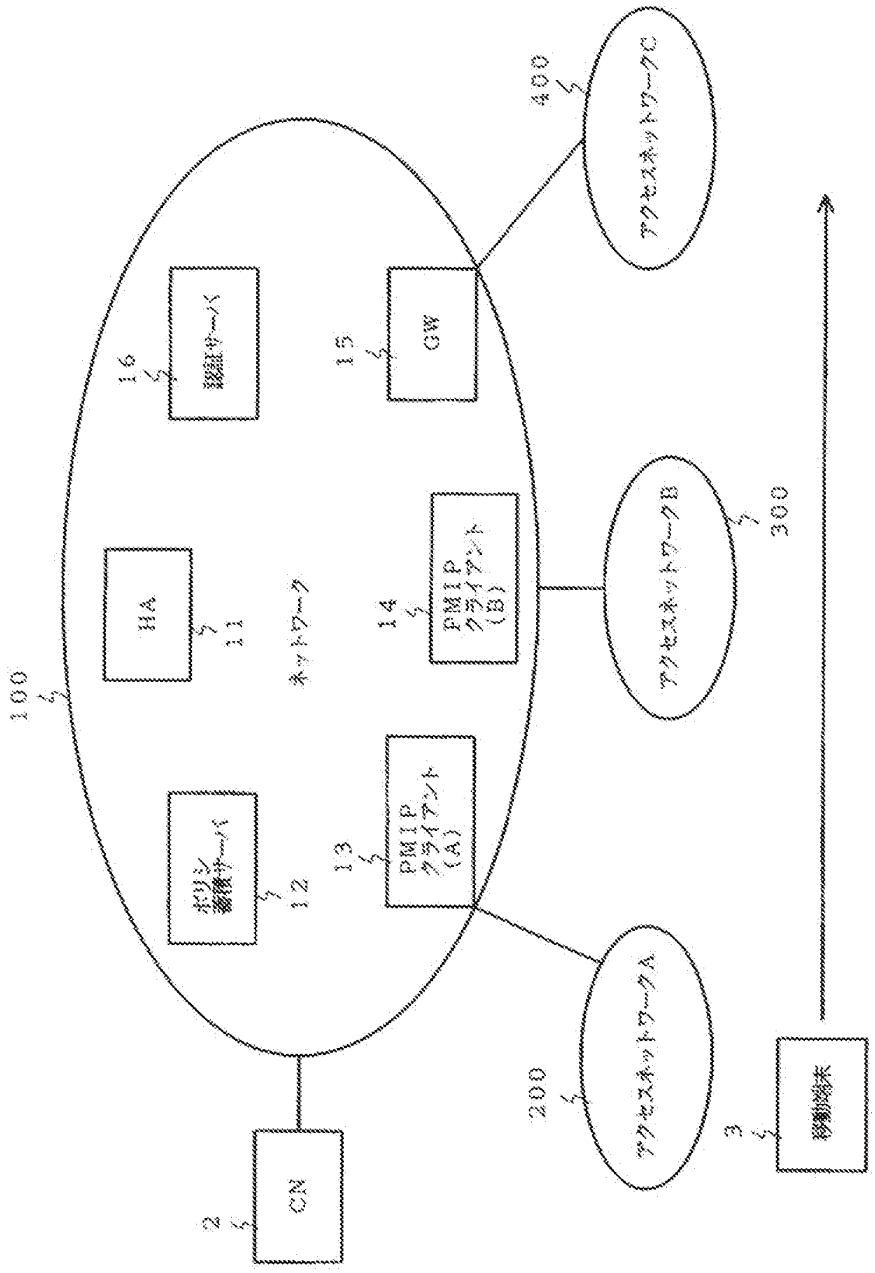
[24] 前記ネットワーク側と前記移動端末との間の信号に、前記通知をやり取りしないCMIP端末の動作を制御するための情報を格納することを特徴とする請求項23記載の移動端末動作制御方法。

[25] 前記プロトコルは、MIPv4(Mobile Internet Protocol version 4)を使用する場合に、AS(Agent Solicitation)、AA(Agent Advertisement)であることを特徴とする請求項23または請求項24記載の移動端末動作制御方法。

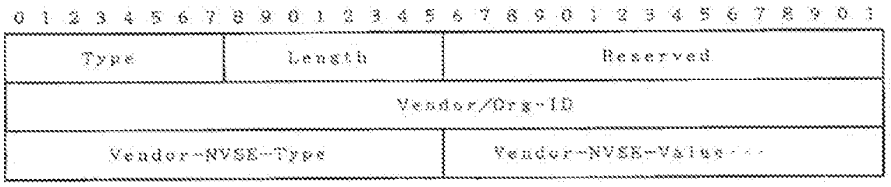
[26] 前記プロトコルは、MIPv6(Mobile Internet Protocol version 7)を使用する場合に、RS(Router Solicitation)、RA(Router Advertisement)であることを特徴とする請求項23または請求項24記載の移動端末動作制御方法。

[27] 前記ネットワーク機器は、3GPP－A(3GPP－Anchor)、MME(Mobility Management Entity)／UPE(User Plane Entity)、PDG(Packet Data Gateway)、PDSN(Packet Data Serving Node)、PDIF(Packet Data Interworking Function)、ASN(Access Service Network)－GW(GateWay)のいずれかであることを特徴とする請求項19から請求項26のいずれか記載の移動端末動作制御方法。

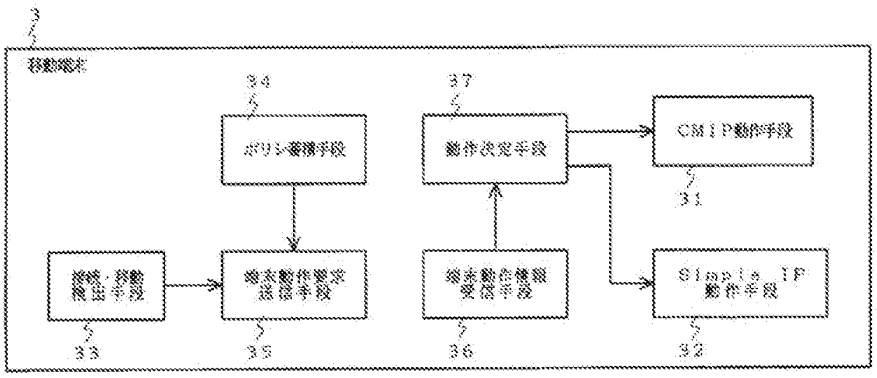
[図1]


100
ネットワーク
HA
11
ホーム
認証サーバ
12
PMIP
クライアント
(A)
13
PMIP
クライアント
(B)
14
GW
15
認証サーバ
16
CN
2
アクセスネットワークA
200
アクセスネットワークB
300
アクセスネットワークC
400
移動端末
3

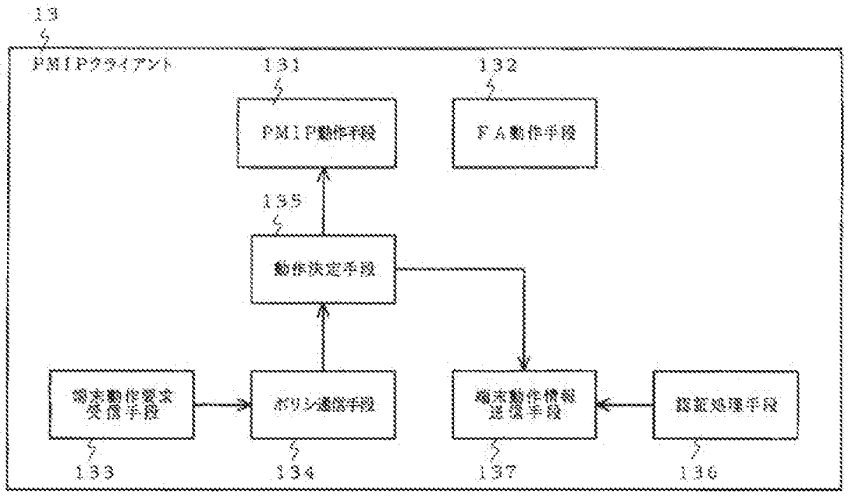
[図2]


0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
Type
Length
Reserved
Vendor/Org-ID
Vendor-NVSE-Type
Vendor-NVSE-Value・・・

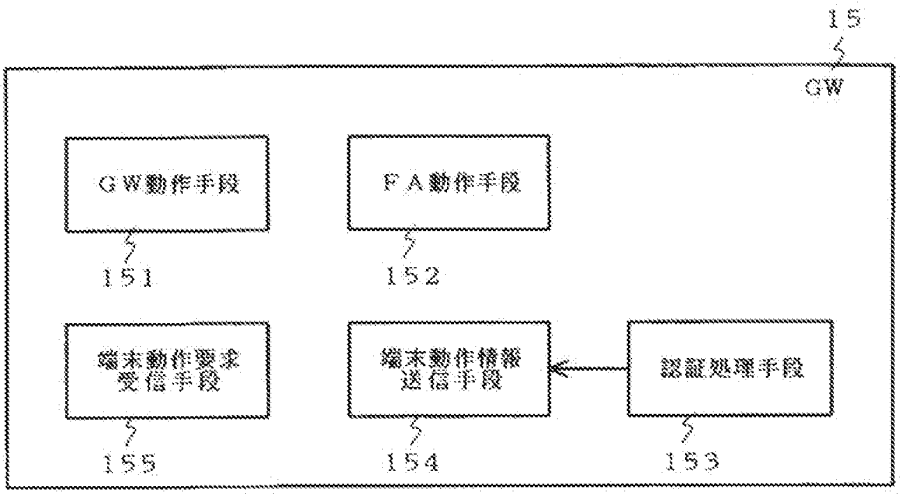
[図3]


3
34
ポリシー管理手段
37
動作決定手段
CMIP動作手段
31
端末動作要求送信手段
端末動作情報受信手段
Simple IP動作手段
33
35
36
32

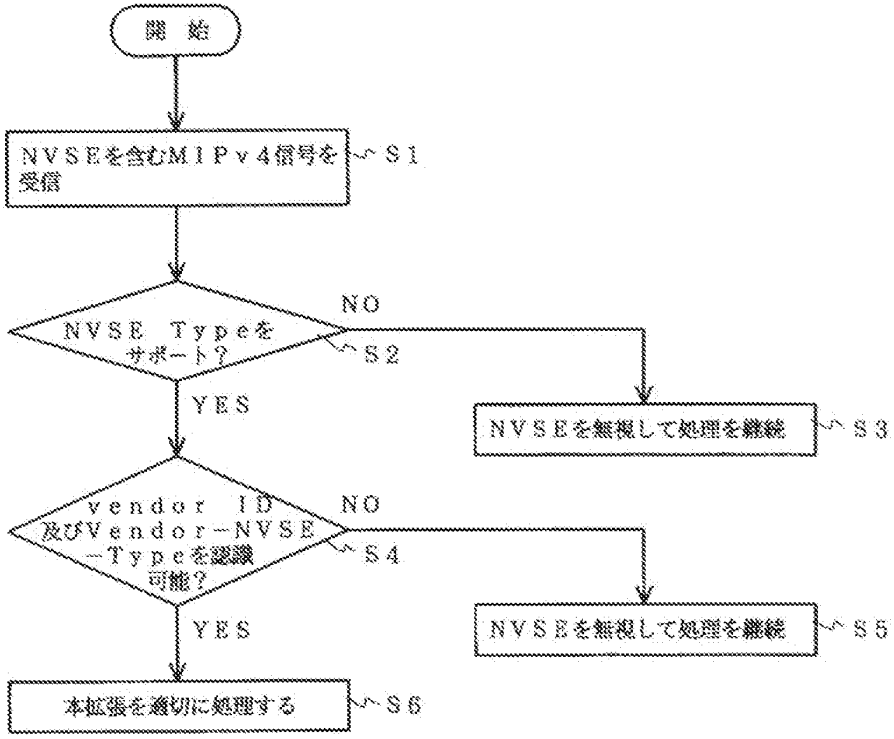
[図4]


13
PMIPクライアント
131
132
PMIP動作手段
FA動作手段
135
動作決定手段
端末動作要求受信手段
ポリシー通信手段
端末動作情報送信手段
認証処理手段
133
134
137
136

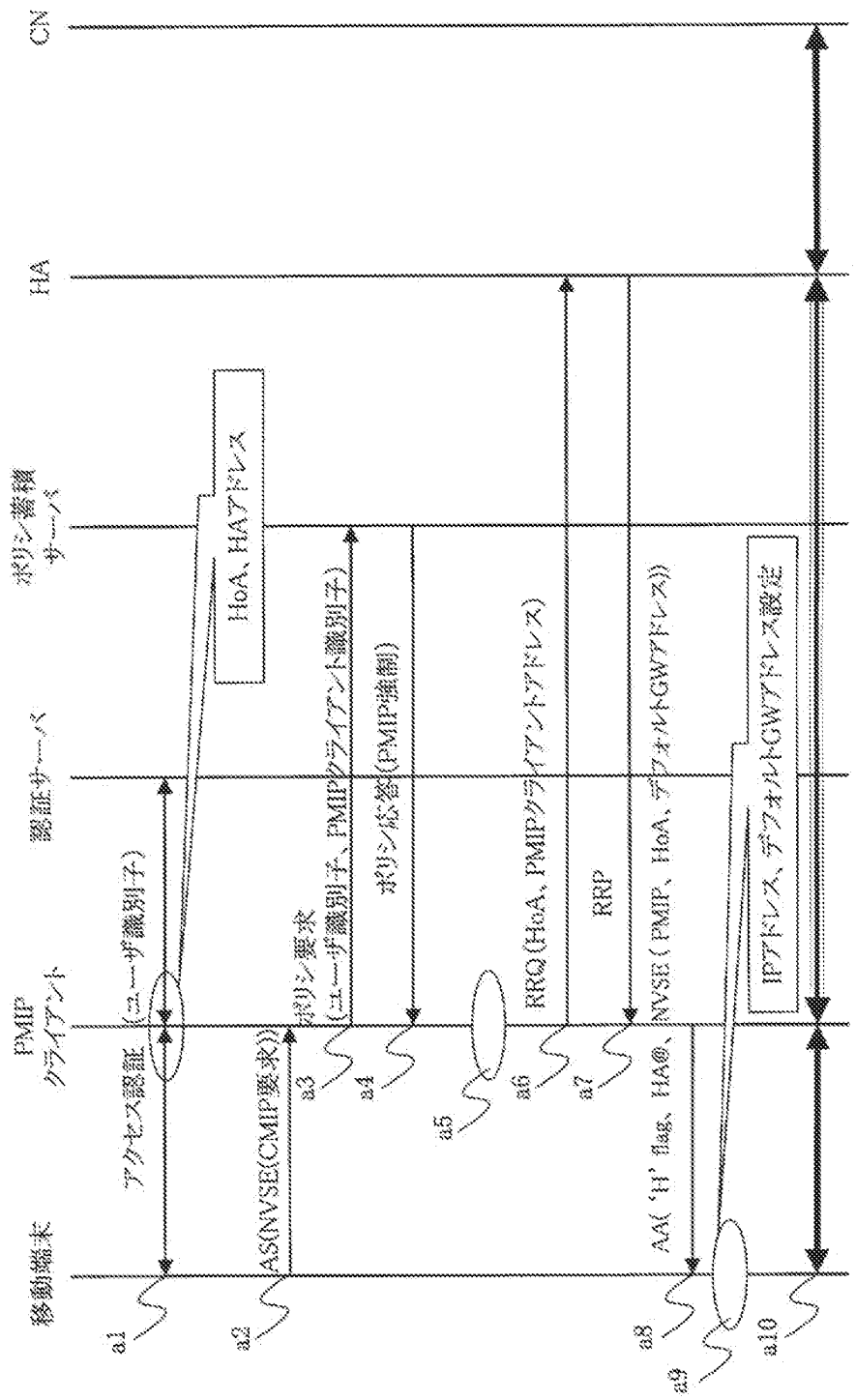
[図5]


15
GW
GW動作手段
FA動作手段
151
152
端末動作要求受信手段
端末動作情報送信手段
認証処理手段
155
154
153

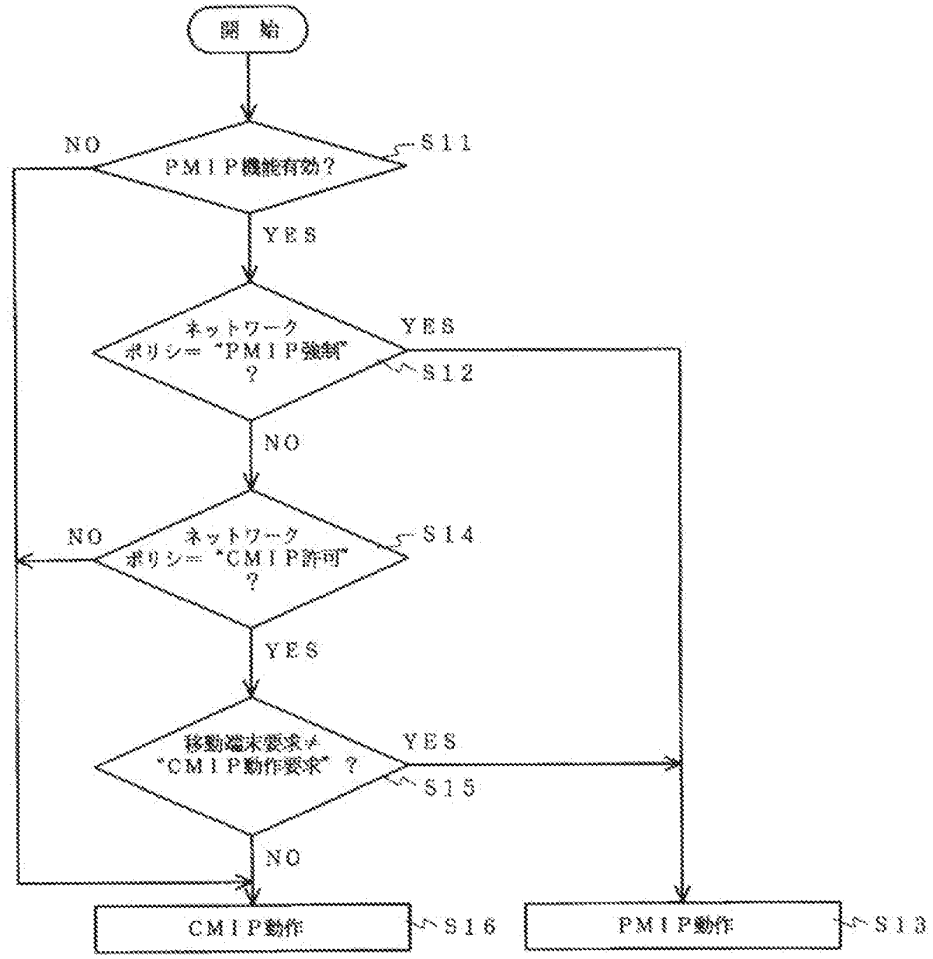
[図6]


開 始
NVSEを含むMIPv4信号を受信
S1
NVSE Typeをサポート?
NO
S2
YES
NVSEを無視して処理を継続
S3
Vendor ID及びVendor−NVSE−Typeを認識可能?
NO
S4
YES
NVSEを無視して処理を継続
S5
本拡張を適切に処理する
S6

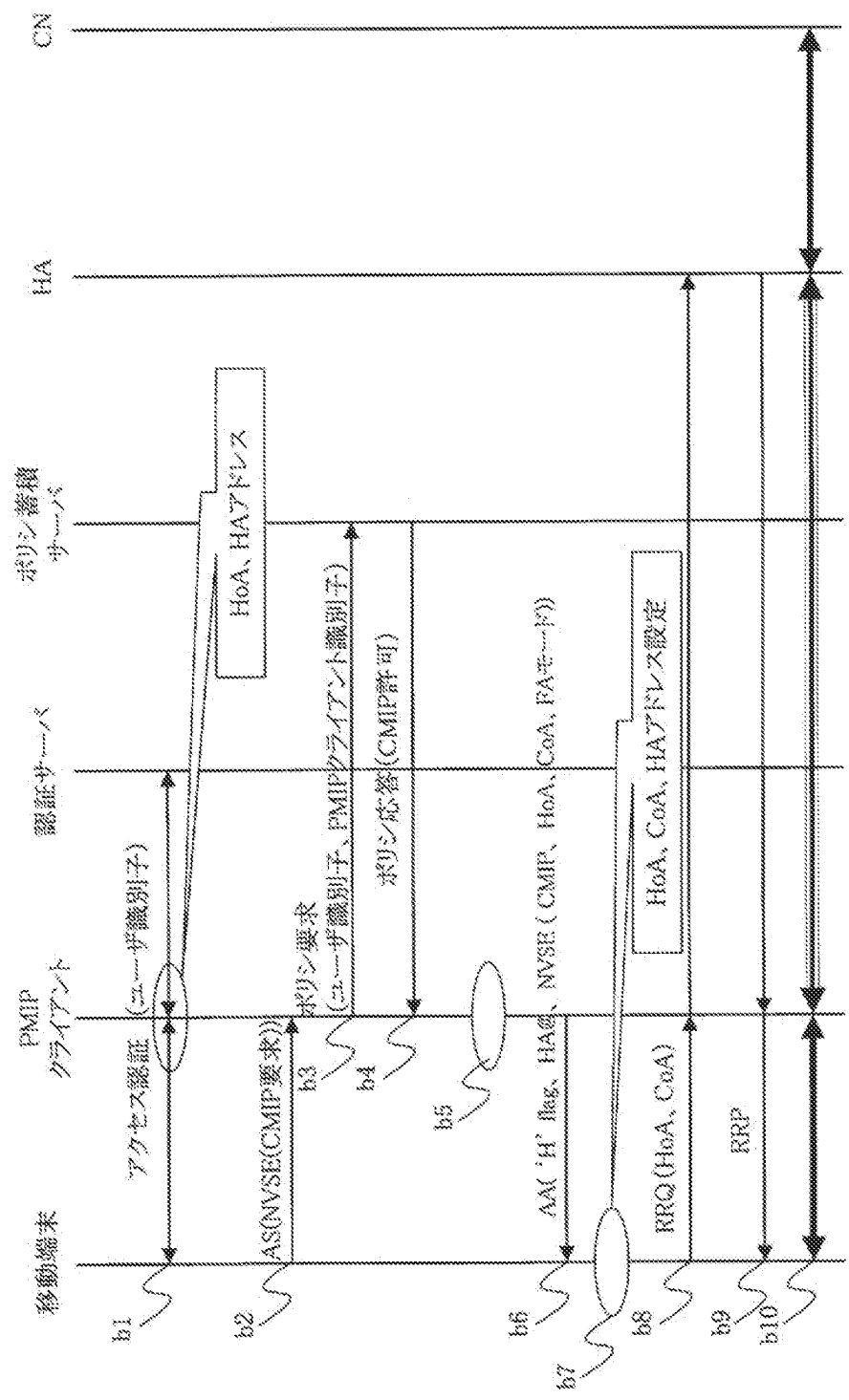
[図7]


CN
HA
ポリシ蓄積サーバ
認証サーバ
PMIPクライアント
移動端末
HoA、HAアドレス
アクセス認証
(ユーザ識別子)
AS(NVSE(CMIP要求))
ポリシ要求
(ユーザ識別子、PMIPクライアント識別子)
ポリシ応答(PMIP強制)
RRQ(HoA、PMIPクライアントアドレス)
RRP
AA('H' flag、HA@、NVSE(PMIP、HoA、デフォルトGWアドレス))
IPアドレス、デフォルトGWアドレス設定
a1
a2
a3
a4
a5
a6
a7
a8
a9
a10

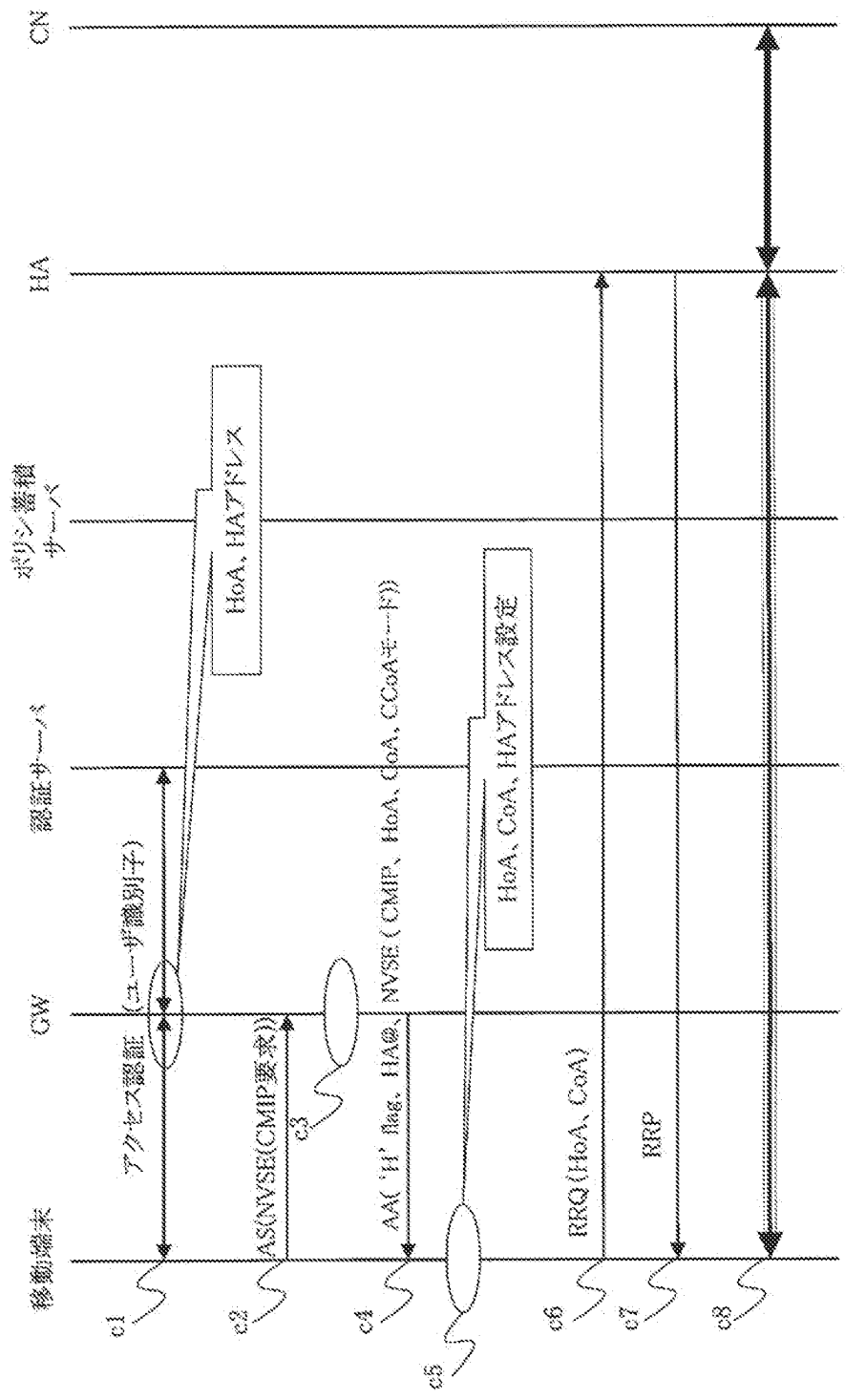
[図8]


開 始
PMIP機能有効?
S11
NO
YES
ネットワーク
ポリシー "PMIP強制"
?
YES
S12
NO
ネットワーク
ポリシー "CMIP許可"
?
S14
NO
YES
移動端末要求が
"CMIP動作要求" ?
YES
S15
NO
CMIP動作
S16
PMIP動作
S13

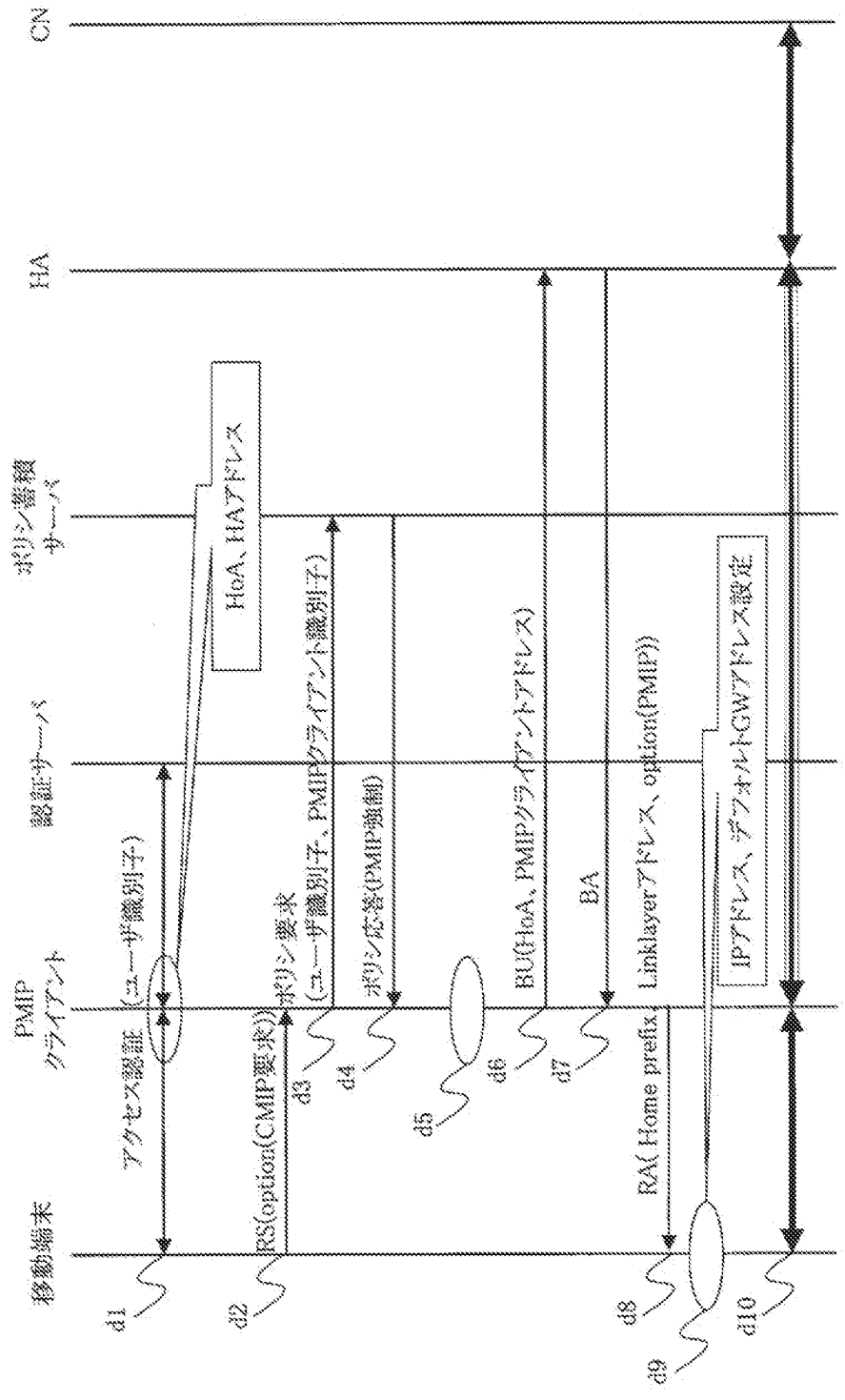
[図9]


CN
HA
ポリシ蓄積サーバ
認証サーバ
PMIPクライアント
移動端末
HoA、HAアドレス
アクセス認証
(ユーザ識別子)
AS(NVSE(CMIP要求))
ポリシ要求
(ユーザ識別子、PMIPクライアント識別子)
ポリシ応答(CMIP許可)
AA('H' flag、HA情、NVSE(CMIP、HoA、CoA、FAキー))
HoA、CoA、HAアドレス設定
RRQ(HoA、CoA)
RRP
b1
b2
b3
b4
b5
b6
b7
b8
b9
b10

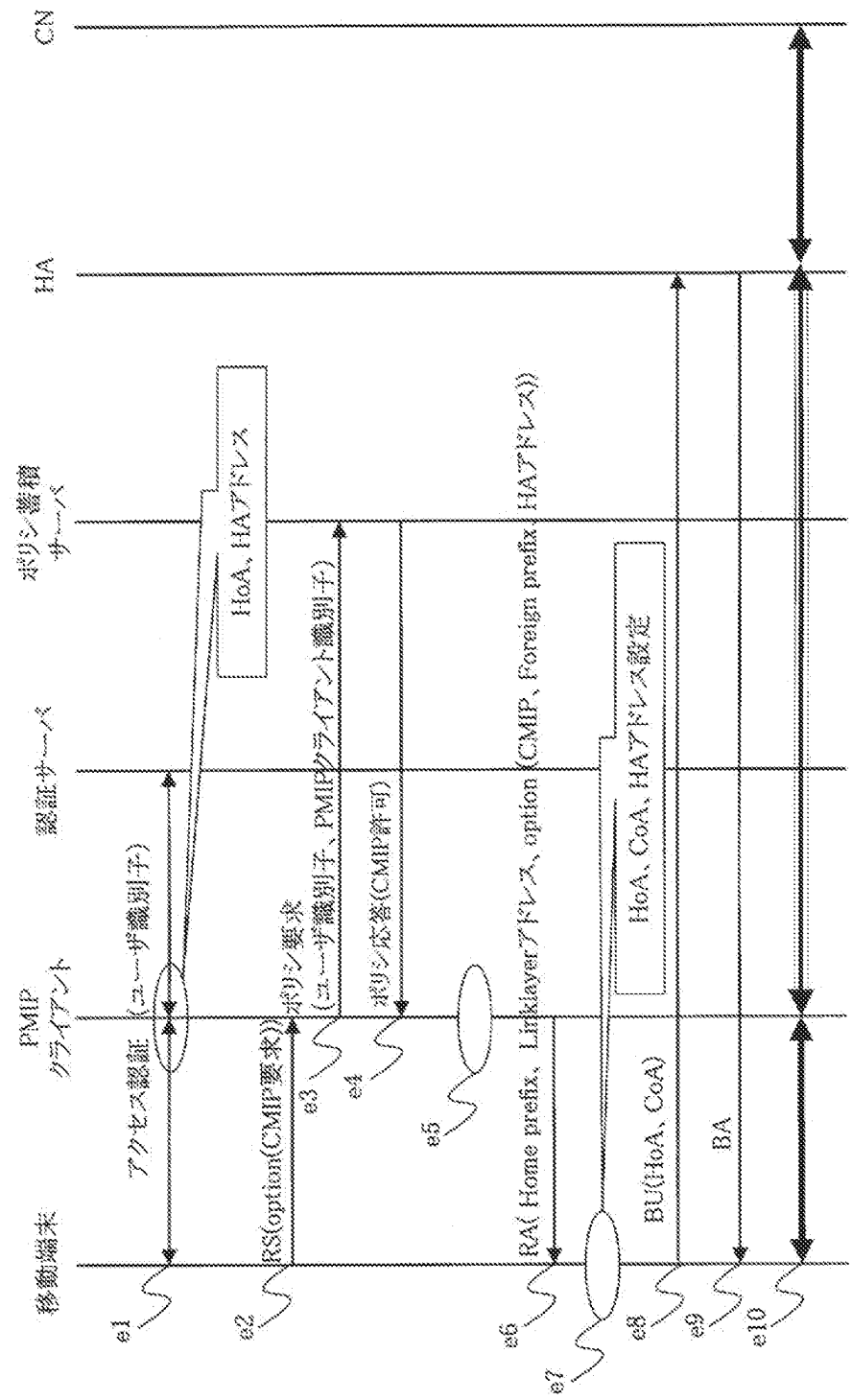
[図10]


CN
HA
ポリシー蓄積サーバ
認証サーバ
GW
移動端末
HoA、HAアドレス
アクセス認証（ユーザ識別子）
AS(NVSE(CMIP要求))
AA('H' flag、HAの、
NVSE (CMIP、HoA、CoA、CCoAモード))
HoA、CoA、HAアドレス設定
RRQ(HoA、CoA)
RRP
c1
c2
c3
c4
c5
c6
c7
c8

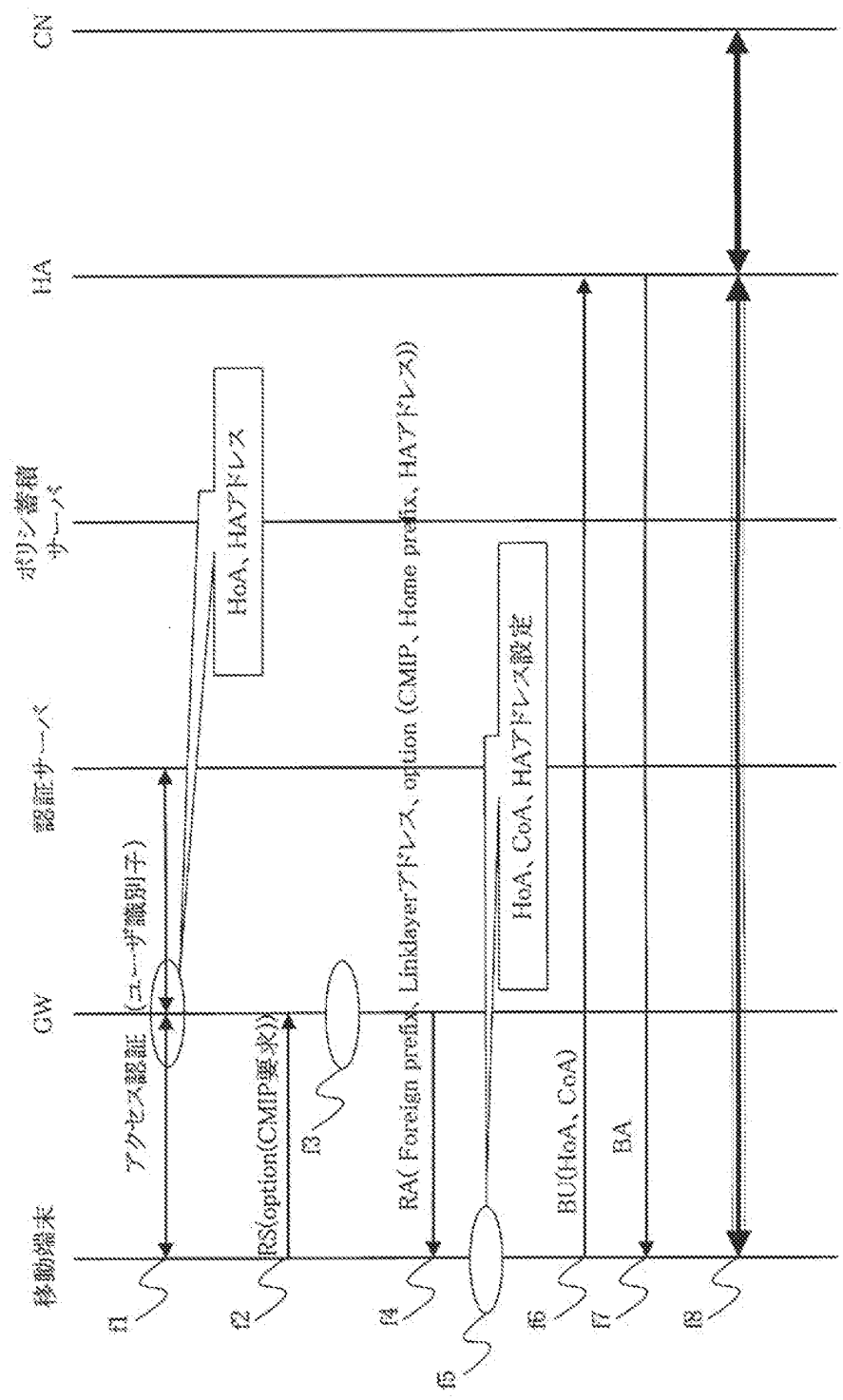
[図11]


CN
HA
ポリシー管理サーバ
認証サーバ
PMIPクライアント
移動端末
HoA、HAアドレス
アクセス認証
(ユーザ識別子)
RS(option(CMIP要求))
ポリシー要求
(ユーザ識別子、PMIPクライアント識別子)
ポリシー応答(PMIP強制)
BU(HoA、PMIPクライアントアドレス)
BA
RA(Home prefix
Linklayerアドレス、option(PMIP))
IPアドレス、デフォルトGWアドレス設定
d1
d2
d3
d4
d5
d6
d7
d8
d9
d10

[図12]


CN
HA
ポリシ蓄積サーバ
認証サーバ
PMIPクライアント
移動端末
アクセス認証
(ユーザ識別子)
HoA、HAアドレス
RS(option(CMIP要求))
ポリシ要求
(ユーザ識別子、PMIPクライアント識別子)
ポリシ応答(CMIP許可)
RA(Home prefix、Linklayerアドレス、option (CMIP、Foreign prefix HAアドレス))
HoA、CoA、HAアドレス設定
BU((HoA、CoA)
BA
e1
e2
e3
e4
e5
e6
e7
e8
e9
e10

[図13]


CN
HA
ポリシ蓄積サーバ
認証サーバ
GW
移動端末
アクセス認証（ユーザ識別子）
HoA、HAアドレス
RS(option(CMIP要求))
RA(Foreign prefix、Linklayerアドレス、option (CMIP、Home prefix、HAアドレス))
HoA、CoA、HAアドレス設定
BU(HoA、CoA)
BA
f1
f2
f3
f4
f5
f6
f7
f8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2008/052236

A. CLASSIFICATION OF SUBJECT MATTER
H04L12/56(2006.01)i, H04Q7/38(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L12/56, H04Q7/38

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2008
Kokai Jitsuyo Shinan Koho 1971-2008 Toroku Jitsuyo Shinan Koho 1994-2008

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	S. Gundavelli et al., Proxy Mobile IPv6, draft-sgundave-mip6-proxymip6-01, IETF NETLMM WG, 2007.01.05, URL<http://www.watersprings.org/pub/id/draft-sgundave-mip6-proxymip6-01.txt> "6.9. Coexistence of CMIP & PMIP Nodes"	1-27
Y	JP 2005-160053 A (Matsushita Electric Industrial Co., Ltd.), 16 June, 2005 (16.06.05), Claims 16, 17; Par. Nos. [0056] to [0068] & EP 1603287 A1 & US 2006/0184692 A1 & WO 2005/039125 A1 & KR 6095729 A & CN 1788471 A	1-27

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 April, 2008 (22.04.08)	01 May, 2008 (01.05.08)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (April 2007)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2008/052236

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2006-166456 A (Microsoft Corp.), 22 June, 2006 (22.06.06), Par. No. [0062] & KR 6063655 A & US 2006/0123128 A1 & EP 1667403 A2	5-8,14-17, 23-26
Y	WO 2005/125118 A1 (Matsushita Electric Industrial Co., Ltd.), 29 December, 2005 (29.12.05), Page 27, line 30 to page 28, line 11 & EP 1610502 A1 & JP 2008-503945	5-8,14-17, 23-26

Form PCT/ISA/210 (continuation of second sheet) (April 2007)

国際調査報告

国際出願番号 PCT/JP2008/052236

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. H04L12/56(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. H04L12/56

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922－1996年
日本国公開実用新案公報 1971－2008年
日本国実用新案登録公報 1996－2008年
日本国登録実用新案公報 1994－2008年

国際調査で使用した電子データベース（データベースの名称，調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Y	S. Gundavelli et al., Proxy Mobile IPv6, draft-sgundave-mip6-proxymip6-01, IETF NETLMM WG, 2007.01.05, URL<http://www.waterspring.org/pub/id/draft-sgundave-mip6-proxymip6-01.txt> "6.9. Coexistence of CMIP & PMIP Nodes"	1-27

☑ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

＊ 引用文献のカテゴリー
「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
22.04.2008	01.05.2008

国際調査機関の名称及びあて先	特許庁審査官（権限のある職員）	5X	3665
日本国特許庁（ISA/JP） 郵便番号100－8915 東京都千代田区霞が関三丁目4番3号	矢頭 尚之 電話番号 03－3581－1101 内線 3596		

様式PCT/ISA/210（第2ページ）（2007年4月）

C（続き）　関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Y	JP 2005-160053 A (松下電器産業株式会社) 2005.06.16, 請求項１６，１７および段落【００５６】－【００６８】 & EP 1603287 A1 & US 2006/0184692 A1 & WO 2005/039125 A1 & KR 6095729 A & CN 1788471 A	1-27
Y	JP 2006-166456 A (マイクロソフトコーポレーション) 2006.06.22, 段落【００６２】 & KR 6063655 A & US 2006/0123128 A1 & EP 1667403 A2	5-8,14-17, 23-26
Y	WO 2005/125118 A1 (松下電器産業株式会社) 2005.12.29, 第２７ページ第３０行－第２８ページ第１１行 & EP 1610502 A1 & JP 2008-503945	5-8,14-17, 23-26

様式PCT/ISA/210（第2ページの続き）（２００７年４月）

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 January 2007 (04.01.2007)

PCT

(10) International Publication Number
WO 2007/001833 A2

(51) International Patent Classification:
H04L 9/32 (2006.01)

(21) International Application Number:
PCT/US2006/023005

(22) International Filing Date: 13 June 2006 (13.06.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/169,093 28 June 2005 (28.06.2005) US

(71) Applicant *(for all designated States except US)*: **SYMBOL TECHNOLOGIES, INC.** [US/US]; One Symbol Plaza, Holtsville, NY 11742 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LINSLEY-HOOD, Richard** [GB/GB]; 5 Park Lane, Bershire, Bracknell RG42 4UL (GB). **D'AGOSTINO, Anthony** [US/US]; 57 Ridge Road, Smithtown, NY 11787 (US). **FAITH, Michael** [US/US]; 11 Woodrige Lane, Coram, NY 11727 (US). **LA ROSA, Dean, V.** [US/US]; 21 Westwind Drive, Bohemia, NY 11716 (US). **PERETZ, Eyal** [US/US]; 3 Singingwood Drive, Stony Brook, NY 11790 (US). **STILL, Gary** [US/US]; 44 Sherman Avenue, Plainview, NY 11803 (US). **UBRIACO, Charles** [US/US]; 1 Chris Court, Northport, NY 11768 (US). **ZANCOLA, Ronald** [US/US]; 2760 Lindenmere Drive, Merrick, NY 11566 (US).

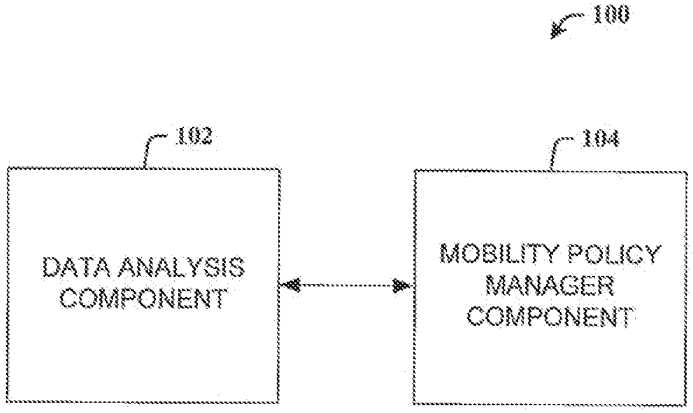
(74) Agents: **AMIN, Himanshu, S.** et al.; Amin & Turocy, LLP, 1900 E. 9th Street, 24th Floor, National City Center, Cleveland, OH 44114 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,

[Continued on next page]

(54) Title: MOBILITY POLICY MANAGER FOR MOBILE COMPUTING DEVICES



(57) Abstract: A mobile computing device that dynamically and automatically manages all network adapters on the device to obtain the best and most effective balance of network connectivity and power consumption without requiring operator or user application intervention. Performance is increased by lowering power consumption (longer battery life) and decreasing transaction latencies, providing a device that has flexibility and ease of use. According to one aspect is a system that includes a data analysis component that monitors parameters associated with a device and a mobility policy manager component that provides balance of network connectivity and power consumption based at least in part upon the monitored stimuli.

WO 2007/001833 A2

RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Title: MOBILITY POLICY MANAGER FOR MOBILE COMPUTING DEVICES

TECHNICAL FIELD

[0001] The following description relates generally to computer systems, and more specifically to balancing network connectivity and managing power consumption of mobile computing devices.

BACKGROUND OF THE INVENTION

[0002] The usage of mobile communication systems utilizing mobile devices has become widespread. Mobile devices such as telephones, pagers, personal digital assistants (PDAs), data terminals, *etc.* are designed to be carried by those who travel from place to place. Each mobile device is capable of communicating by wireless communication between the mobile device and an access point to which the mobile device is registered. As the mobile device roams from one cell or geographic region to another, the mobile device will typically deregister with the access point of the previous cell and register with the access point associated with the new cell.

[0003] In general, when one network component transmits a signal to another network component, the transmitting network component typically expects a response from the receiving network component within a reasonable time frame. If a response is not received within the reasonable time (*e.g.*, lapsed time-out), communication usually terminates. In order to respond to such transmissions in a timely manner, many devices continuously operate in a high power (*e.g.*, full power) mode such that a device can continuously, periodically and/or on demand, monitor, receive and respond to incoming transmissions.

[0004] Current industry trend, however, is to minimize power consumption. Common techniques that mitigate power consumption include automatic power management utilities that transition a device to a lower power or "off" state. For example, many devices utilize standard power management technologies, such as Advanced Configuration and Power Interface (ACPI), which enable system software/firmware (*e.g.*, operating system or application) to control power by automatically transitioning the device to a low power state (*e.g.*, Standby, Suspend, Hibernate, Sleep, Deep Sleep…) when full power is not required. Typically, when the device transitions to the lower power state, network connections are terminated and power is removed from the CPU, volatile memory, transceivers, *etc.* The

foregoing can conserve power consumption and battery life, however, the device essentially is inoperable to communicate with the outside world until it transitions back to a higher power state.

[0005] Mobile computing devices lack the ability to configure and/or automatically and dynamically manage power consumption of multiple network adapters and cost of, or options for, use of network connectivity. This places a burden upon a user to either statically predict and set or manually configure the nominal operating mode periodically. This does not facilitate flexibility and/or adaptability during operation of the device. Therefore, what is needed is an optimum power-connectivity trade-off that can be continuously obtained and updated automatically and/or dynamically on a device.

SUMMARY OF THE INVENTION

[0006] The following presents a simplified summary of the invention in order to provide a basic understanding of some aspects of the invention. This summary is not an extensive overview of the invention. It is intended to neither identify key or critical elements of the invention nor delineate the scope of the invention. Its sole purpose is to present some concepts of the invention in a simplified form as a prelude to the more detailed description that is presented later.

[0007] The invention disclosed and claimed herein, in one aspect thereof, comprises a portable and/or mobile communication device that includes a data analysis component and a mobility policy manager component. The data analysis component monitors stimuli associated with a device. The stimuli can be at least one of network traffic, an internet protocol address, and a port number. The mobility policy manager (MPM) component provides balancing of network connectivity and power consumption based at least in part upon the monitored stimuli. The MPM component can include a policy module that determines a policy based at least in part on the monitored stimuli. The MPM component can include an alter module that alters at least one policy to provide an optimal device power consumption. Further included in the MPM component is a tolerance parameter module that ensures the policy remains within a predetermined power consumption range and/or a hysteresis module that prevents excessive altering between multiple device states. According to a further aspect, the MPM component can include a persistence module that allows the device to remain in a non-optimal state during a communication. Also included

can be a route module that hinders traffic in the absence of an appropriate route. The route module can include a firewall module that restricts traffic if an appropriate route is not available.

[0008] Still another aspect is a communication architecture including a network that includes a plurality of resources, a data analysis component that captures data about the plurality of resources, and a mobility policy manager (MPM) component. The MPM component can determine a policy to obtain efficient network connectivity and power consumption based at least in part on the captured data. The captured data includes at least one statistic based on network traffic associated with an Internet Protocol address or a port number end point. The MPM component can dynamically alter a predefined policy based at least in part upon the captured data.

[0009] According to a further aspect is a method for dynamically managing network connection performance. The method includes creating an initial policy, monitoring the initial policy, and automatically altering the initial policy to achieve at least one policy objective. The initial policy is created based upon at least one statistic derived from device conditions. Automatically altering the initial policy includes determining a state of a connection, ascertaining if the connection is complete, and maintaining the connection state until communication is complete.

[0010] Yet another aspect is a computer readable medium of a portable and/or mobile device to efficiently manage network connections. The computer readable medium includes a data analysis component operable to monitor device conditions and a mobility policy manager component that dynamically alters a policy to achieve at least one policy objective. The mobility policy manager provides balance of network connectivity and power consumption based at least in part on the monitored device conditions.

[0011] The following description and the annexed drawings set forth in detail certain illustrative aspects of the invention. These aspects are indicative, however, of but a few of the various ways in which the principles of the invention may be employed and the invention is intended to include all such aspects and their equivalents. Other advantages and novel features of the invention will become apparent from the following detailed description of the invention when considered in conjunction with the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1 illustrates a system for managing network connectivity and power consumption.

[0013] FIG. 2 illustrates exemplary modules associated with a system that facilitates managing power consumption and network connectivity.

[0014] FIG. 3 illustrates exemplary modules associated with a system that facilitates managing power consumption and network connectivity.

[0015] FIG. 4 illustrates modules associated with a system that facilitates managing power consumption and network connectivity.

[0016] FIG. 5 illustrates a system for managing power consumption utilizing an artificial intelligence component.

[0017] FIG. 6 illustrates a flow chart of a methodology for optimal network connectivity and minimum power consumption utilizing one or more policies based on external and/or internal device conditions.

[0018] FIG. 7 illustrates a flow chart of a methodology for dynamically and automatically managing power consumption and enhancing network connectivity

[0019] FIG. 8 illustrates a flow chart of a methodology for determining transition of power states between a plurality of adapters.

[0020] FIG. 9 illustrates a schematic block diagram of a computer operable to execute the disclosed systems and/or methods.

[0021] FIG. 10 illustrates an exemplary device operative to execute the systems and/or methods disclosed herein.

[0022] FIG. 11 illustrates an exemplary networking environment that can be employed in connection with the invention.

[0023] FIG. 12 illustrates an exemplary operating environment that can be employed in connection with the invention.

DESCRIPTION OF THE INVENTION

[0024] The subject invention is now described with reference to the drawings, wherein like reference numerals are used to refer to like elements throughout. In the following description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the invention. It may be evident, however, that the invention may be practiced without these specific details.

In other instances, well-known structures and devices are shown in block diagram form in order to facilitate describing the invention.

[0025] As used in this application, the terms "component," "module" and "system" are intended to refer to a computer-related entity, either hardware, a combination of hardware and software, software, or software in execution. For example, a component may be, but is not limited to being, a process running on a processor, a processor, an object, an executable, a thread of execution, a program, and/or a computer. By way of illustration, both an application running on a server and the server can be a component. One or more components may reside within a process and/or thread of execution and a component may be localized on one computer and/or distributed between two or more computers.

[0026] The word "exemplary" is used herein to mean serving as an example, instance, or illustration. Any aspect or design described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other aspects or designs.

[0027] Furthermore, the invention may be implemented as a method, apparatus, or article of manufacture using standard programming and/or engineering techniques to produce software, firmware, hardware, or any combination thereof to control a computer to implement the disclosed invention. The term "article of manufacture" as used herein is intended to encompass a computer program accessible from any computer-readable device, carrier, or media. For example, computer readable media can include but are not limited to magnetic storage devices (*e.g.*, hard disk, floppy disk, magnetic strips…), optical disks (*e.g.*, compact disk (CD), digital versatile disk (DVD)…), smart cards, and flash memory devices (*e.g.*, card, stick). Additionally it should be appreciated that a carrier wave can be employed to carry computer-readable electronic data such as those used in transmitting and receiving electronic mail or in accessing a network such as the Internet or a local area network (LAN). Of course, those skilled in the art will recognize many modifications may be made to this configuration without departing from the scope or spirit of the subject invention.

[0028] Artificial intelligence based systems (*e.g.*, explicitly and/or implicitly trained classifiers) can be employed in connection with performing inference and/or probabilistic determinations and/or statistical-based determinations as in accordance with one or more aspects of the subject invention as described hereinafter. As used

herein, the terms "infer" and "inference" refer generally to the process of reasoning about or inferring states of the system, environment, and/or user from a set of observations as captured *via* events and/or data. Inference can be employed to identify a specific context or action, or can generate a probability distribution over states, for example. The inference can be probabilistic - that is, the computation of a probability distribution over states of interest based on a consideration of data and events. Inference can also refer to techniques employed for composing higher-level events from a set of events and/or data. Such inference results in the construction of new events or actions from a set of observed events and/or stored event data, whether or not the events are correlated in close temporal proximity, and whether the events and data come from one or several event and data sources. Various classification schemes and/or systems (*e.g.*, support vector machines, neural networks, expert systems, Bayesian belief networks, fuzzy logic, data fusion engines…) can be employed in connection with performing automatic and/or inferred action in connection with the subject invention.

[0029] Referring initially to FIG. 1, illustrated is a system 100 for managing network connectivity and power consumption. The system 100 can be implemented on any mobile device, such as a personal digital assistant (PDA), a mobile phone, a smart phone, a palm pilot, a personal data assistant, a mobile computer, a scanner, an RFID device, a laptop (or desktop) computer, a wearable computer device or any other mobile computing device.

[0030] The system 100 includes a data analysis component 102 that interfaces with a mobility policy manager component 104. The data analysis component 102 is adapted to receive, obtain, monitor, record, *etc.* stimuli and/or data (both internal and external to the device) that relates to connectivity and power consumption of the device. The data analysis component 102 is further adapted to analyze and evaluate such data and can provide statistics generated from such data. For example, the data analysis component 102 can obtain information relating to network traffic to and/or from an Internet Protocol address and/or port number end points or end point pairs. Port numbers allow different applications on the same device to utilize network resources without interfering with each other. The data analysis component 102 can further monitor power consumption to ensure that, through a policy, the power consumption remains within a predetermined range.

[0031] The data analysis component 102 interfaces with the mobility policy manager component 104 to allow a mobile computing device to dynamically and automatically manage all wide-area network (WAN), local-area network (LAN), personal area network (PAN), cradle, modem or other network adapters on a device. This facilitates an efficient and optimum balance of network connectivity and power consumption. In such a way, operator and/or user application interaction is mitigated after initial device configuration.

[0032] The mobility policy manager 104 receives the analyzed data from the data analysis component 102, such as in the form of statistics. The statistics are utilized to either create a predetermined fixed policy and/or to allow the mobility policy manager 104 to automatically evaluate the cost versus the performance requirements of the system to create and/or dynamically alter and maintain any such policy. The policy or set of policies are generated for use on the portable and/or mobile communication device and identify what actions should and/or should not be taken for a given set of stimuli. The policy, written in software, can be generated for use in a particular mobile communication environment. The policy can treat physical network adapters that allow multiple power settings internally as multiple separate logical adapters.

[0033] With reference now to FIG. 2, illustrated are modules associated with a system 200 that facilitates managing power consumption and network connectivity of a mobile computing device. The modules illustrated are exemplary and it is to be understood that more or less modules can be utilized and fall within the scope of the specification and appended claims. The mobility policy manager component 104 can include a policy module 202, an alter module 204 and/or a parameter module 206 that interface with each other and with the external data analysis component 102.

[0034] The policy module 202 can receive information, such as from the external data analysis component 102 and, based upon derived statistics, determine a policy to achieve an optimal balance of network connectivity and power consumption. The statistics can be collected ahead of time to allow for creating an initial policy. Thus, the policy can be predetermined and fixed, allowing no changes or alterations to such policy. Alternatively or in addition, the policy can be dynamic allowing the device to automatically evaluate the cost versus performance requirements of the device. The policy can be created and/or dynamically altered to maintain a policy that facilitates network connectivity and reduces power consumption.

[0035] An alter module 204 allows the policy to be dynamically updated based upon changing conditions (both internal and external) as received from the data analysis component 102. The alter module 204 utilizes the statistics, for example, and may limit or force particular modifications to a policy depending on the situation, such as allowing minimum and maximum outcomes to be predetermined through an interface with a parameter module 206. Thus, the policy can be dynamically altered to ensure that the policy is within an expected range or set of circumstances that provide the device with optimum connectivity and reduced power consumption.

[0036] FIG. 3 illustrates exemplary modules associated with a system 300 that facilitates managing power consumption and network connectivity. The exemplary modules illustrated can be utilized in conjunction with the one or more modules discussed with reference to FIG. 2 or they can be utilized separately. Included are a packet route management module 302, hysteresis module 304, duration module 306, persistence module 308, and/or route module 310. The mobility policy manager component 104 utilizes various parameters of a device to ensure that power consumption is reduced while allowing optimum network connectivity and device functionality. The mobility policy manager component 104 can specify an initial policy to establish and/or maintain the connectivity and/or functionality objectives and further can dynamically and/or autonomously update, modify, *etc.* the policy to ensure the objectives are met. This can be accomplished through the one or more modules associated with the system 300.

[0037] The packet route management module 302 is adapted to allow more than one network adapter associated with a device to be powered or active at any given time, such as a substantially similar time. In such a situation, the packet route management module 302 can analyze the various network adapters and routes and alter a policy to control the connectivity of the device. For example, each network adapter may be analyzed separately or concurrently with the other adapters to determine needed policy changes, if any, to ensure that each network adapter has optimum connectivity. The packet route management module 302 can be further adapted to stop a connection, either permanently or temporarily, if a particular connection is interfering with one or more other connections, or if a particular connection is experiencing problems, such as too much traffic on the network, slow network performance, unresponsive network connections, *etc.*

[0038] Another module associated with the mobility policy manger component 104 is a hysteresis module 304 that analyzes one or more network connections and determines the state of each connection. For example, the state of a network connection can be described as various levels of power consumption. These levels can include "off," "listening for network availability," "connected to a network," "receiving data", "transmitting data," *etc.* It is to be understood that each of these levels can include addition levels relating to, for example, the geographic distance between the terminal and an infrastructure transmitter depending on the network type. The levels can also be described as being hierarchically ordered, for example, from the least power consumption ("off") to the highest power consumption ("transmitting data"). The hysteresis module 304 is further adapted to prevent inefficient, continuous alternating between multiple states and may force the device into a particular state to improve device efficiency.

[0039] The duration module 306 is adapted to measure, obtain, determine, *etc.* the time interval/range spent at each of the power levels for any given connection. The time interval/range can include the amount of time taken to get to a particular state and the dwell time to transition out of such state. The time interval/range can also be measured as including the time spent in a particular state. Thus, the duration module 306 can construct a "total duty cycle" of power consumption for a given connection based on the time spent at a given power level and the power used at that level. This can be incorporated into the profile and/or statistic(s) where appropriate and any subsequent attempt to connect can be matched against such existing profile(s) to determine the most power effective manner to complete that connection. It should be understood that the measurement does not have to be "time" provided there is a measurement means that can determine the range in which a given connection is at a particular power level.

[0040] The hysteresis module 304 and the duration module 306 provide the mobility policy manager component 104 information to determine whether to switch to the next lower power usage adapter and/or whether to switch to a higher power usage adapter. The capacity to determine if it is valid to transition from a lower power state to the next higher one, or vice versa, is part of the consideration of the overall switching characteristics of the mobility policy manager component 104. The decision to switch to a next lower (or higher) power usage adapter can be determined by leaving it in a particular state and determining its availability at that level before,

moving it to a lower (or higher) state. A higher power usage adapter can then be transitioned to a lower power consumption state possible after some dwell time to allow for a degree of hysteresis.

[0041] A persistence module 308 is adapted to monitor one or more network connections and can work separately or in conjunction with the hysteresis module 304. The persistence module 308 can monitor the one or more network connections. If a particular connection is not complete (*e.g.*, data, information, messages, text is still being transferred to/from the device) the persistence module 308 can force the device to remain in a particular state for the duration of that particular connection. For example, it may be determined that the state of a particular connection is non-optimal. By non-optimal it is meant that the connection is inefficient, consumes excessive resources, is not the best configuration, *etc.* The persistence module 308 analyzes the connection and can determine that, although in a non-optimal state, the particular state should be maintained for the duration of that particular connection or connections. This allows for the correct completion of that particular communication. After the communication is complete, the persistence module 308 can allow the device to switch to a more optimal state.

[0042] A route module 310 is provided that can determine a network connection route, alter an existing route, force a particular route, and/or block or reject traffic if no appropriate route is available. Exemplary modules associated with the route module 310 are illustrated in FIG. 4 and can include an override module 402, a categorization module 404, and/or a firewall module 406. The exemplary modules 402-406 interface with each other and the other modules of the mobility policy manager 104 to provide systems and/or methods that balance network connectivity and manage power consumption of mobile devices.

[0043] The override module 402 provides input to the mobility policy manager 104 to alter the policy to allow for designated connections based upon one or more policy goal(s) and/or objective(s). The designated connections can be determined based upon statistics relating to such network connections. A determination can be made to designate a particular route or outcome by overriding other choices as determined by analysis of the statistics and/or policy. The override module 402 facilitates forcing the particular outcome or route though utilization of the designated connections. For example, if there is more than one adapter on a device, each having a different degree of power consumption, the mobility policy manager

component 104 is adapted to determine which adapter to utilize and can transition between the adapters, depending on the requirements of a particular connection. For example, if it is desired to download a large amount of data onto a device, the mobility policy manager component 104 can determine that speed is more important than power conservation (since more power generally equals more speed). In such a situation, the device will transition to the highest power/speed adapter, regardless of any policy power conservation requirements and/or objectives.

[0044] The connections can also be categorized though utilization of a categorization module 404. Such a categorization can analyze the requirements of creating a new route and/or maintaining an existing route. The categorization module 404 can prefer one requirement over the other. For example, the categorization module 404 can treat the requirements for creating a new route differently than the requirements of utilizing an existing route. In some situations, a new route can be preferred, while in other situations the existing route may be preferred. The categorization module 404 is adapted to analyze the requirements of both in relation to the policy goals, objectives, statistic information, and/or information obtained from other modules of the mobility policy component 104 and/or data analysis component 102.

[0045] There are some circumstances in which particular traffic may be blocked and/or rejected though utilization of a firewall module 406. If a determination is made that no appropriate route for a network connection exists, such a connection may be refused and/or hindered and connectivity not allowed. In this situation, the connection will need to be made or attempted at a different time when an appropriate route is available.

[0046] FIG. 5 illustrates a system 500 for managing power consumption that employs artificial intelligence (AI) which facilitates automating one or more features in accordance with the subject invention. The AI can be effected *via* AI component 502, as illustrated, that interfaces with an external data analysis component 102 and a mobility policy manager component 104.

[0047] The subject invention (*e.g.*, in connection with managing power consumption) can employ various AI-based schemes for carrying out various aspects thereof. For example, a process for determining the best and most effective balance of network connectivity and power consumption can be facilitated *via* an automatic classifier system and process. Moreover, where network adapters are employed

having the same or similar resources, the classifier can be employed to determine which network adapter to employ in a particular situation.

[0048] A classifier is a function that maps an input attribute vector, x = (x1, x2, x3, x4, *xn*), to a confidence that the input belongs to a class, that is, f(x) = *confidence*(*class*). Such classification can employ a probabilistic and/or statistical-based analysis (*e.g.*, factoring into the analysis utilities and costs) to prognose or infer an action that a user desires to be automatically performed.

[0049] A support vector machine (SVM) is an example of a classifier that can be employed. The SVM operates by finding a hypersurface in the space of possible inputs, which hypersurface attempts to split the triggering criteria from the non-triggering events. Intuitively, this makes the classification correct for testing data that is near, but not identical to training data. Other directed and undirected model classification approaches include, *e.g.*, naïve Bayes, Bayesian networks, decision trees, neural networks, fuzzy logic models, and probabilistic classification models providing different patterns of independence can be employed. Classification as used herein also is inclusive of statistical regression that is utilized to develop models of priority.

[0050] As will be readily appreciated from the subject specification, the subject invention can employ classifiers that are explicitly trained (*e.g.*, *via* a generic training data) as well as implicitly trained (*e.g.*, *via* observing user behavior, receiving extrinsic information). For example, SVM's are configured *via* a learning or training phase within a classifier constructor and feature selection module. Thus, the classifier(s) can be used to automatically learn and perform a number of functions, including but not limited to determining according to a predetermined criteria when to grant access, which stored procedure to execute, which policy objective/goal to conform to or override, *etc.* The criteria can include, but is not limited to, the amount of data or resources to accessed *via* a call, the type of data, the importance of the data, *etc.*

[0051] In view of the exemplary systems shown and described above, methodologies, which may be implemented in accordance with one or more aspects of the invention, will be better appreciated with reference to the diagram of FIGS. 6, 7, and 8. The systems and/or methodologies automatically and dynamically increase performance of a device while lowering power consumption (longer battery life) and decreasing transaction latencies, providing ease-of-use and device flexibility. While,

for purposes of simplicity of explanation, the methodologies are shown and described as a series of function blocks, it is to be understood and appreciated that the invention is not limited by the order of the blocks, as some blocks may, in accordance with the invention, occur in different orders and/or concurrently with other blocks from that shown and described herein. Moreover, not all illustrated blocks may be required to implement a methodology in accordance with one or more aspects of the invention. It is to be appreciated that the various blocks may be implemented *via* software, hardware, a combination thereof or any other suitable means (*e.g.* device, system, process, component) for carrying out the functionality associated with the blocks. It is also to be appreciated that the blocks are merely to illustrate certain aspects of the invention in a simplified form and that these aspects may be illustrated *via* a lesser and/or greater number of blocks.

[0052] FIG. 6 illustrates a flow chart of a methodology 600 for optimal network connectivity and minimum power consumption utilizing one or more policy based on external and/or internal device conditions. The method starts, at 602, where a policy is determined. The policy can be a fixed policy or a configurable policy and can be based on statistics derived prior to the policy being implemented. The statistics can be derived from network traffic to and/or from an Internet Protocol address and/or port number end points or end point pairs. The statistics can alternatively or in addition be derived from other stimuli, information, data, *etc.* derived from both an internal and/or an external environment of the device. For example, the stimuli can be information derived from a plurality of resources in a communication architecture. The plurality of resources can be other devices, data being communicated, network adapters, *etc.*

[0053] The statistics can be monitored and/or automatically evaluated, at 604, such that the policy can be altered to allow minimum and maximum outcomes to be pre-determined. The policy can be evaluated to ensure that the connectivity of a particular connection is at an optimal level, and if not at an optimal level to determine if the connection should be maintained based on other factors, such as a continuing communication. An objective of the policy can be to automatically evaluate the cost versus the performance requirements to create and/or dynamically alter and maintain any such policy.

[0054] At 606, the policy is altered, maintained, created, *etc.* based upon the evaluation performed. For example, the policy can be altered to control the

connectivity of the device, maintain the current device connectivity, *etc.* The policy can be modified to determine a particular network connection route, alter an existing route, force a particular route, and/or block or reject traffic if no appropriate route is available.

[0055] Referring now to FIG. 7, illustrated is a flow chart of a methodology 700 for dynamically and automatically managing power consumption of multiple network adapters and cost of, or options for, use of network connectivity. The method starts, at 702, where external and/or internal stimuli is gathered and, based on such stimuli, statistics are created. The statistics can relate to connectivity and/or power consumption of a mobile computing device. For example, the data can be related to network traffic to and/or from an Internet Protocol address and/or port number end points or end point pairs. The network traffic data can be utilized to derive statistics regarding network connectivity and the optimal connections, routes, options, *etc.* to both achieve efficient network communication and reduce costs associated with operating and maintaining a mobile computing device.

[0056] A policy is created based upon the derived statistics, at 704. The policy can be a fixed policy that does not allow changes or an adaptable policy that allows automatic evaluation of cost versus performance requirements. An automatic evaluation allows dynamic creation, alteration, maintenance, *etc.* of the policy to ensure that network connectivity and power consumption is at optimum levels.

[0057] Data relating to the operation, connectivity, power consumption, *etc.* of a mobile device is periodically and/or constantly analyzed, at 706, to ensure the objectives established by the one or more policies is being met. The performance of the device and its network connectivity is evaluated, at 708, based upon the analyzed data. The policy can be altered, at 710, if necessary to conform to the policy requirements. For example, the policy can be altered to allow for persistence of a non-optimal state for the duration of a particular connection or connections to allow for the correct completion of communication. The policy can allow for appropriate hysteresis to prevent inefficient or continuous alternating between multiple states that would otherwise result in inefficient communication. The policy can allow designated connections that override other choices to force a particular outcome or route. The policy can treat physical network adapters that allow multiple power settings internally as multiple separate logical adapters. The policy can be modified or adapted to establish and achieve objective relating to device connectivity and

reduced power consumption. The policy can further be utilized to perform other functions relating to conserving power while maintaining optimum network connectivity and data transfer.

[0058] With reference now to FIG. 8, illustrated is a flow chart of a methodology 800 for determining transition of power states between a plurality of radio or network adapters. For illustration purposes, the methodology will be discussed with reference to three network adapters ("A," "B," and "C"). It is to be understood that this methodology can be applied to two or more network adapters and is not to be limited by the following detailed description.

[0059] The method starts, at 802, with receipt of network adapter power requirements. The power requirements of each adapter can be ranked from the adapter with the highest power usage ("A") to the adapter with the lowest power usage ("C"). Methods of distinguishing between the states of each adapter other than a ranking can be utilized. It is also to be understood that adapters may have similar power usages and can be classified together.

[0060] At 804, the state of each adapter is received. The state can refer to the power level consumption and can be categorized in ranges consisting of an hierarchical order, such as 1 being the lowest power consumption and 5 being the highest power consumption. For example, the states can be "off" = 1; "listening for network availability" = 2; "connected to a network" = 3; "receiving data" = 4; and "transmitting data" = 5. It is to be understood that more or less power level consumption states can be utilized as well as each level having additional levels included therein. For example, the additional levels can relate to the geographic distance between the terminal and the infrastructure transmitter depending on the network type.

[0061] The method continues, at 806, where the availability of a particular adapter to transition to a lower or higher state is received. With continuing reference to the example above, adapter "A" consumes the most power and adapter "C" the least power. Thus, it would be advantageous to keep adapter "A" at the lowest power consumption state possible and adapter "C" in the highest power consumption state, if required. For example, the adapters can have the following initial state: adapter "A" = 1 (off); adapter "B" = 3 (connected to a network); and adapter "C" = 2 (listening for network availability). If adapter "C" indicates that it is capable of transiting to state 3, then the option to switch the state of adapter "C" becomes valid and adapter "C" can

change states with an adapter with a higher power consumption rate. Adapter "A", having the highest power consumption, would change to state 3 only if the adapter(s) having the lower power consumption (adapters "B" and "C" in this illustration) indicate it is no longer in coverage and is now unusable.

[0062] At 808, the adapters are changed to the appropriate state to comply with optimal power consumption and network connectivity as determined by policy goals and/or objective. For example, when adapter "C" transitions to state 3, adapter "B" can, in turn, transition to a lower state, such as state 1 or 2. By way of example, the initial states of the adapters could be "A" is in state 3; adapter "B" is in state 2; and adapter "C" is in state 1. Thus, the method would determine that the adapter with the highest power consumption ("A") is in the highest power state (3) and thus, if adapter "B" and/or adapter "C" is available, such adapter should transition to state 3 and adapter "A" transition to state 2 or lower. In this situation, the ideal transition would be for adapter "A" to transition to state 1 and adapter "C" to transition to state 3. However, policy objectives and/or goals might not allow such minimum levels of power consumption.

[0063] Referring now to FIG. 9, illustrated is a schematic block diagram of a portable and/or mobile hand-held terminal device 900 (similar to the portable scanning device 1000 as illustrated in FIG. 10) according to one aspect of the invention, in which a processor 902 is responsible for controlling the general operation of the device 900. The processor 902 is programmed to control and operate the various components within the device 900 in order to carry out the various functions described herein. The processor 902 can be any of a plurality of suitable processors. The manner in which the processor 902 can be programmed to carry out the functions relating to the invention will be readily apparent to those having ordinary skill in the art based on the description provided herein.

[0064] A memory 904 connected to the processor 902 serves to store program code executed by the processor 902, and serves as a storage means for storing information such as user credential and receipt transaction information and the like. The memory 904 can be a non-volatile memory suitably adapted to store at least a complete set of the information that is displayed. Thus, the memory 904 can include a RAM or flash memory for high-speed access by the processor 902 and/or a mass storage memory, *e.g.*, a micro drive capable of storing gigabytes of data that comprises text, images, audio, and video content. According to one aspect, the

memory 904 has sufficient storage capacity to store multiple sets of information, and the processor 902 could include a program for alternating or cycling between various sets of display information.

[0065] A display 906 is coupled to the processor 902 *via* a display driver system 908. The display 906 can be a color liquid crystal display (LCD), plasma display, or the like. In this example, the display 906 is a ¼ VGA display with sixteen levels of gray scale. The display 906 functions to present data, graphics, or other information content. For example, the display 906 can display a set of customer information, which is displayed to the operator and can be transmitted over a system backbone (not shown). Additionally, the display 906 can display a variety of functions that control the execution of the device 900. The display 906 is capable of displaying both alphanumeric and graphical characters.

[0066] Power is provided to the processor 902 and other components forming the hand-held device 900 by an onboard power system 910 (*e.g.*, a battery pack). In the event that the power system 910 fails or becomes disconnected from the device 900, a supplemental power source 912 can be employed to provide power to the processor 902 and to charge the onboard power system 910. The processor 902 of the device 900 induces a sleep mode to reduce the current draw upon detection of an anticipated power failure.

[0067] The terminal 900 includes a communication subsystem 914 that includes a data communication port 916, which is employed to interface the processor 802 with a remote computer. The port 916 can include at least one of Universal Serial Bus (USB) and IEEE 1394 serial communications capabilities. Other technologies can also be included, for example, infrared communication utilizing an infrared data port.

[0068] The device 900 can also include a radio frequency (RF) transceiver section 918 in operative communication with the processor 902. The RF section 918 includes an RF receiver 920, which receives RF signals from a remote device *via* an antenna 922 and demodulates the signal to obtain digital information modulated therein. The RF section 918 also includes an RF transmitter 924 for transmitting information to a remote device, for example, in response to manual user input *via* a user input device 926 (*e.g.*, a keypad) or automatically in response to the completion of a transaction or other predetermined and programmed criteria. The transceiver section 918 facilitates communication with a transponder system, for example, either

passive or active, that is in use with product or item RF tags. The processor 902 signals (or pulses) the remote transponder system *via* the transceiver 918, and detects the return signal in order to read the contents of the tag memory. In one implementation, the RF section 818 further facilitates telephone communications using the device 900. In furtherance thereof, an audio I/O section 928 is provided as controlled by the processor 902 to process voice input from a microphone (or similar audio input device) and audio output signals (from a speaker or similar audio output device). In another implementation, the device 900 can provide voice recognition capabilities such that when the device 900 is used simply as a voice recorder, the processor 902 can facilitate high-speed conversion of the voice signals into text content for local editing and review, and/or later download to a remote system, such as a computer word processor. Similarly, the converted voice signals can be used to control the device 900 instead of using manual entry *via* the keypad 926.

[0069] Onboard peripheral devices, such as a printer 930, signature pad 932, and a magnetic strip reader 934 can also be provided within the housing of the device 900 or accommodated externally through one or more of the external port interfaces 916.

[0070] The device 900 can also include an image capture system 936 such that the user can record images and/or short movies for storage by the device 900 and presentation by the display 906. Additionally, a dataform reading system 938 is included for scanning dataforms. It is to be appreciated that these imaging systems (936 and 938) can be a single system capable of performing both functions.

[0071] FIG. 10 is provided to assist in understanding and to provide context to an embodiment of the invention. Specifically, FIG. 10 illustrates an example of a handheld terminal 1000 operative to execute the systems and/or methods disclosed herein.

[0072] The handheld terminal 1000 includes a housing 1002 which can be constructed from a high strength plastic, metal, or any other suitable material. The handheld terminal 1000 includes a display 1004. As is conventional, the display 1004 functions to display data or other information relating to ordinary operation of the handheld terminal 1000 and/or mobile companion (not shown). For example, software operating on the handheld terminal 1000 and/or mobile companion can provide for the display of various information requested by the user. Additionally, the display 1004 can display a variety of functions that are executable by the handheld

terminal 1000 and/or one or more mobile companions. The display 1004 provides for graphics based alphanumerical information such as, for example, the price of an item requested by the user. The display 1004 also provides for the display of graphics such as icons representative of particular menu items, for example. The display 1004 can also be a touch screen, which can employ capacitive, resistive touch, infrared, surface acoustic wave, or grounded acoustic wave technology.

[0073] The handheld terminal 1000 further includes user input keys 1006 for allowing a user to input information and/or operational commands. The user input keys 1006 can include a full alphanumeric keypad, function keys, enter keys, *etc.* The handheld terminal 1000 can also include a magnetic strip reader 1008 or other data capture mechanism (not shown). An electronic signature apparatus can also be employed in connection with the magnetic strip reader or a telecheck system.

[0074] The handheld terminal 1000 can also include a window 1010 in which a bar code reader/bar coding imager is able to read a bar code label, or the like, presented to the handheld terminal 1000. The handheld terminal 1000 can include a light emitting diode (LED) (not shown) that is illuminated to reflect whether the bar code has been properly or improperly read. Alternatively, or additionally, a sound can be emitted from a speaker (not shown) to alert the user that the bar code has been successfully imaged and decoded. The handheld terminal 1000 also includes an antenna (not shown) for wireless communication with a radio frequency (RF) access point; and an infrared (IR) transceiver (not shown) for communication with an IR access point.

[0075] FIG. 11 is a schematic block diagram of a sample-computing environment 1100 with which the invention can interact. The system 1100 includes one or more client(s) 1110. The client(s) 1110 can be hardware and/or software (*e.g.*, threads, processes, computing devices). The system 1100 also includes one or more server(s) 1120. The server(s) 1120 can also be hardware and/or software (*e.g.*, threads, processes, computing devices). The servers 1120 can house threads to perform transformations by employing the invention, for example.

[0076] One possible communication between a client 1110 and a server 1120 can be in the form of a data packet adapted to be transmitted between two or more computer processes. The system 1100 includes a communication framework 1140 that can be employed to facilitate communications between the client(s) 1110 and the server(s) 1120. The client(s) 1110 are operably connected to one or more client data

store(s) 1150 that can be employed to store information local to the client(s) 1110. Similarly, the server(s) 1120 are operably connected to one or more server data store(s) 1130 that can be employed to store information local to the servers 1140.

[0077] With reference to FIG. 12, an exemplary environment 1210 for implementing various aspects of the invention includes a computer 1212. The computer 1212 includes a processing unit 1214, a system memory 1216, and a system bus 1218. The system bus 1218 couples system components including, but not limited to, the system memory 1216 to the processing unit 1214. The processing unit 1214 can be any of various available processors. Dual microprocessors and other multiprocessor architectures also can be employed as the processing unit 1214.

[0078] The system bus 1218 can be any of several types of bus structure(s) including the memory bus or memory controller, a peripheral bus or external bus, and/or a local bus using any variety of available bus architectures including, but not limited to, Industrial Standard Architecture (ISA), Micro-Channel Architecture (MSA), Extended ISA (EISA), Intelligent Drive Electronics (IDE), VESA Local Bus (VLB), Peripheral Component Interconnect (PCI), Card Bus, Universal Serial Bus (USB), Advanced Graphics Port (AGP), Personal Computer Memory Card International Association bus (PCMCIA), Firewire (IEEE 1394), and Small Computer Systems Interface (SCSI).

[0079] The system memory 1216 includes volatile memory 1220 and nonvolatile memory 1222. The basic input/output system (BIOS), containing the basic routines to transfer information between elements within the computer 1212, such as during start-up, is stored in nonvolatile memory 1222. By way of illustration, and not limitation, nonvolatile memory 1222 can include read only memory (ROM), programmable ROM (PROM), electrically programmable ROM (EPROM), electrically erasable ROM (EEPROM), or flash memory. Volatile memory 1220 includes random access memory (RAM), which acts as external cache memory. By way of illustration and not limitation, RAM is available in many forms such as synchronous RAM (SRAM), dynamic RAM (DRAM), synchronous DRAM (SDRAM), double data rate SDRAM (DDR SDRAM), enhanced SDRAM (ESDRAM), Synchlink DRAM (SLDRAM), and direct Rambus RAM (DRRAM).

[0080] Computer 1212 also includes removable/non-removable, volatile/non-volatile computer storage media. FIG. 12 illustrates, for example a disk storage 1224. Disk storage 1224 includes, but is not limited to, devices like a magnetic disk drive,

floppy disk drive, tape drive, Jaz drive, Zip drive, LS-100 drive, flash memory card, or memory stick. In addition, disk storage 1224 can include storage media separately or in combination with other storage media including, but not limited to, an optical disk drive such as a compact disk ROM device (CD-ROM), CD recordable drive (CD-R Drive), CD rewritable drive (CD-RW Drive) or a digital versatile disk ROM drive (DVD-ROM). To facilitate connection of the disk storage devices 1224 to the system bus 1218, a removable or non-removable interface is typically used such as interface 1226.

[0081] It is to be appreciated that FIG. 12 describes software that acts as an intermediary between users and the basic computer resources described in suitable operating environment 1210. Such software includes an operating system 1228. Operating system 1228, which can be stored on disk storage 1224, acts to control and allocate resources of the computer system 1212. System applications 1230 take advantage of the management of resources by operating system 1228 through program modules 1232 and program data 1234 stored either in system memory 1216 or on disk storage 1224. It is to be appreciated that the invention can be implemented with various operating systems or combinations of operating systems.

[0082] A user enters commands or information into the computer 1212 through input device(s) 1236. Input devices 1236 include, but are not limited to, a pointing device such as a mouse, trackball, stylus, touch pad, keyboard, microphone, joystick, game pad, satellite dish, scanner, TV tuner card, digital camera, digital video camera, web camera, and the like. These and other input devices connect to the processing unit 1214 through the system bus 1218 *via* interface port(s) 1238. Interface port(s) 1238 include, for example, a serial port, a parallel port, a game port, and a universal serial bus (USB). Output device(s) 1240 use some of the same type of ports as input device(s) 1236. Thus, for example, a USB port may be used to provide input to computer 1212, and to output information from computer 1212 to an output device 1240. Output adapter 1242 is provided to illustrate that there are some output devices 1240 like monitors, speakers, and printers, among other output devices 1240, which require special adapters. The output adapters 1242 include, by way of illustration and not limitation, video and sound cards that provide a means of connection between the output device 1240 and the system bus 1218. It should be noted that other devices and/or systems of devices provide both input and output capabilities such as remote computer(s) 1244.

[0083] Computer 1212 can operate in a networked environment using logical connections to one or more remote computers, such as remote computer(s) 1244. The remote computer(s) 1244 can be a personal computer, a server, a router, a network PC, a workstation, a microprocessor based appliance, a peer device or other common network node and the like, and typically includes many or all of the elements described relative to computer 1212. For purposes of brevity, only a memory storage device 1246 is illustrated with remote computer(s) 1244. Remote computer(s) 1244 is logically connected to computer 1212 through a network interface 1248 and then physically connected *via* communication connection 1250. Network interface 1248 encompasses communication networks such as local-area networks (LAN) and wide-area networks (WAN). LAN technologies include Fiber Distributed Data Interface (FDDI), Copper Distributed Data Interface (CDDI), Ethernet, Token Ring and the like. WAN technologies include, but are not limited to, point-to-point links, circuit switching networks like Integrated Services Digital Networks (ISDN) and variations thereon, packet switching networks, and Digital Subscriber Lines (DSL).

[0084] Communication connection(s) 1250 refers to the hardware/software employed to connect the network interface 1248 to the bus 1218. While communication connection 1250 is shown for illustrative clarity inside computer 1212, it can also be external to computer 1212. The hardware/software necessary for connection to the network interface 1248 includes, for exemplary purposes only, internal and external technologies such as, modems including regular telephone grade modems, cable modems and DSL modems, ISDN adapters, and Ethernet cards.

[0085] What has been described above includes examples of the invention. It is, of course, not possible to describe every conceivable combination of components or methodologies for purposes of describing the subject invention, but one of ordinary skill in the art may recognize that many further combinations and permutations of the invention are possible. Accordingly, the invention is intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims.

[0086] In particular and in regard to the various functions performed by the above described components, devices, circuits, systems and the like, the terms (including a reference to a "means") used to describe such components are intended to correspond, unless otherwise indicated, to any component which performs the specified function of the described component (*e.g.*, a functional equivalent), even

though not structurally equivalent to the disclosed structure, which performs the function in the herein illustrated exemplary aspects of the invention. In this regard, it will also be recognized that the invention includes a system as well as a computer-readable medium having computer-executable instructions for performing the acts and/or events of the various methods of the invention.

[0087] In addition, while a particular feature of the invention may have been disclosed with respect to only one of several implementations, such feature may be combined with one or more other features of the other implementations as may be desired and advantageous for any given or particular application. Furthermore, to the extent that the terms "includes," and "including" and variants thereof are used in either the detailed description or the claims, these terms are intended to be inclusive in a manner similar to the term "comprising."

CLAIMS

What is claimed is:

1. A portable communication device comprising:

a data analysis component that monitors stimuli associated with a device; and

a mobility policy manager (MPM) component that provides balancing of network connectivity and power consumption based at least in part upon the monitored stimuli.

2. The portable communication device of claim 1, the data analysis component monitors at least one of network traffic, an internet protocol address, and a port number.

3. The portable communication device of claim 1, the MPM component comprising a policy module that determines a policy based at least in part on the monitored stimuli.

4. The portable communication device of claim 1, the MPM component comprising an alter module that alters at least one policy to provide an optimal device power consumption.

5. The portable communication device of claim 4, the MPM component further comprising a tolerance parameter module that ensures the policy remains within a predetermined power consumption range.

6. The portable communication device of claim 1, the MPM component comprising a hysteresis module that prevents excessive altering between multiple device states.

7. The portable communication device of claim 1, the MPM component comprising a persistence module that allows the device to remain in a non-optimal state during a communication.

8. The portable communication device of claim 7, the non-optimal state is maintained until completion of the communication.

9. The portable communication device of claim 1, the MPM component comprising a packet route management module that controls a connectivity of a first network adapter and at least a second network adapter that are powered at a substantially similar time.

10. The portable communication device of claim 1, the MPM component comprising a designate module that designates at least one connection.

11. The portable communication device of claim 11, the designate module overrides a connection to force one of an outcome and a route.

12. The portable communication device of claim 1, the MPM component comprising a route module that hinders traffic in the absence of an appropriate route.

13. The portable communication device of claim 12, the route module comprising a categorization module that categorizes connections that distinguishes between requirements to create a new route and preference of an existing route.

14. The portable communication device of claim 12, the route module comprising a firewall module that restricts traffic if an appropriate route is not available.

15. A communication architecture, comprising:

a network that includes a plurality of resources;

a data analysis component that captures data about the plurality of resources; and a mobility policy manager (MPM) component that determines a policy to obtain efficient network connectivity and power consumption based at least in part on the captured data.

16. The communication architecture of claim 15, the captured data comprising at least one statistic based on network traffic associated with an Internet Protocol address or a port number end point.

17. The communication architecture of claim 15, the MPM component dynamically alters a predefined policy based at least in part upon the captured data.

18. A method for dynamically managing network connection performance, comprising:

creating an initial policy;

monitoring the initial policy; and

automatically altering the initial policy to achieve at least one policy objective.

19. The method of claim 18, the initial policy is created based upon at least one statistic derived from device conditions.

20. The method of claim 18, automatically altering the initial policy further comprising:

determining a state of a connection;

ascertaining if the connection is complete; and

maintaining the connection state until communication is complete.

21. The method of claim 18, automatically altering the initial policy further includes one of allowing, denying, and rejecting traffic.

22. A computer readable medium of a portable device to efficiently manage network connections, comprising:

a data analysis component operable to monitor device conditions; and

a mobility policy manager component that dynamically alters a policy to achieve at least one policy objective.

23. The computer readable medium of claim 22, the mobility policy manager provides balancing of network connectivity and power consumption based at least in part on the monitored device conditions.

24. The computer readable medium of claim 22, an initial policy is created based at least in part on the monitored device conditions.

25. A portable communication system, comprising:

means for monitoring device connectivity and power consumption;

means for creating a policy based at least in part on the monitored device connectivity and power consumption; and

means for altering the policy based on monitored operating conditions.

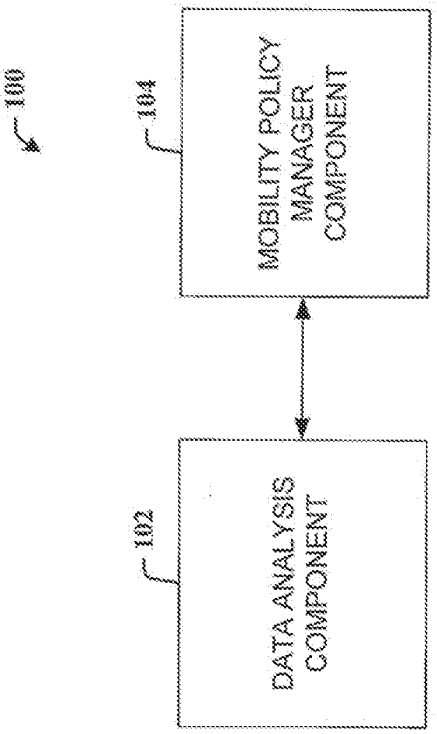

100
104
MOBILITY POLICY MANAGER COMPONENT
102
DATA ANALYSIS COMPONENT

FIG. 1

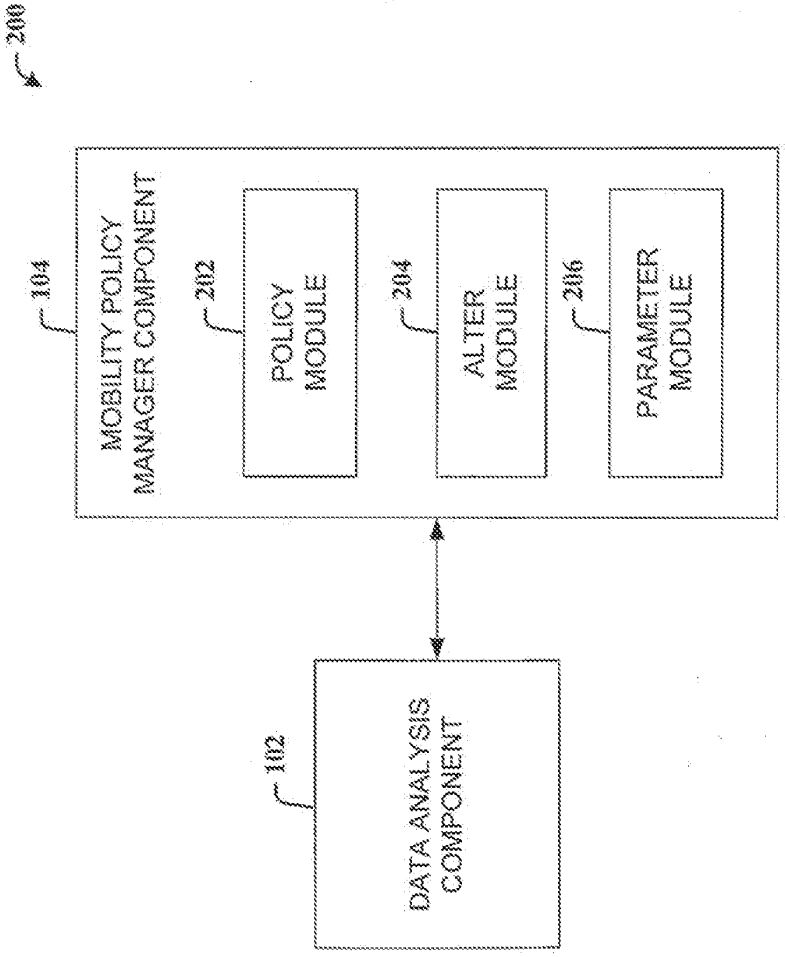

200
104
MOBILITY POLICY MANAGER COMPONENT
202
POLICY MODULE
204
ALTER MODULE
206
PARAMETER MODULE
102
DATA ANALYSIS COMPONENT

FIG. 2

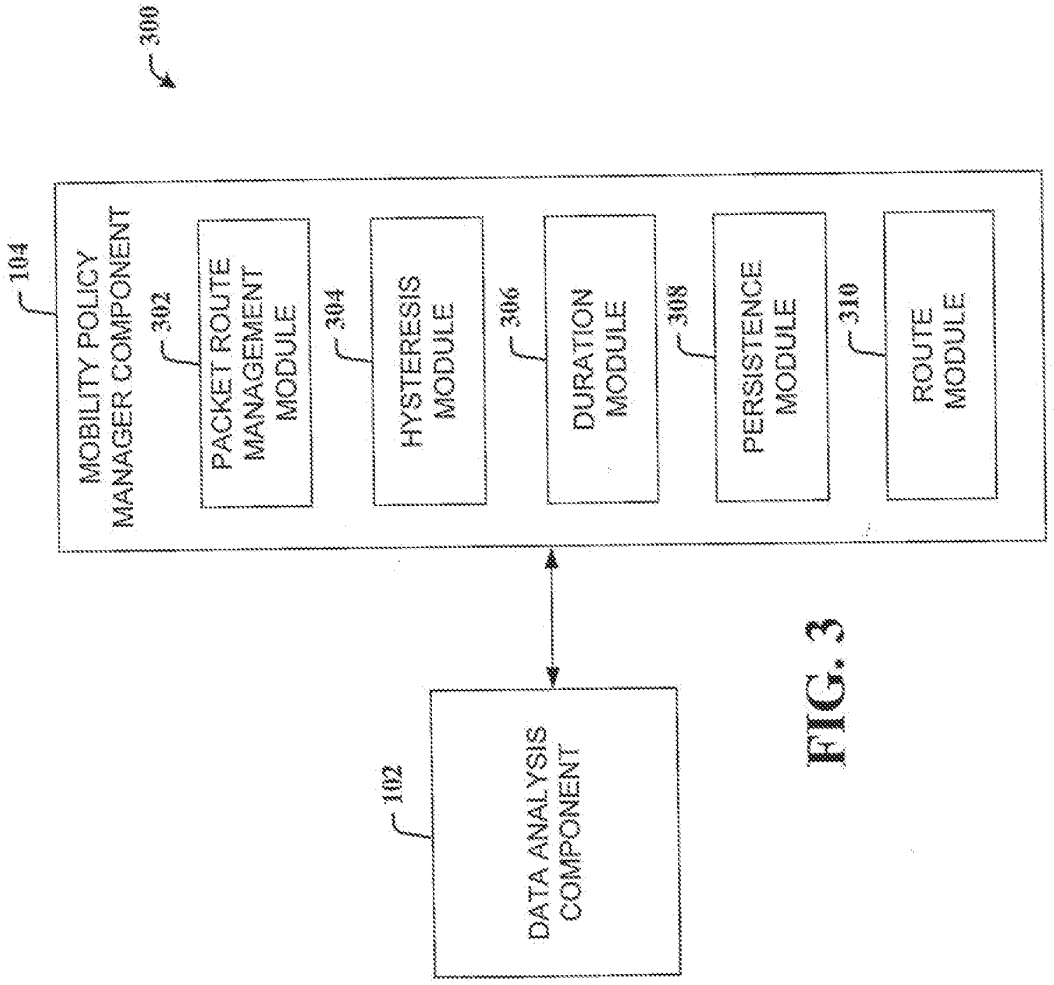

300
104
MOBILITY POLICY MANAGER COMPONENT
302
PACKET ROUTE MANAGEMENT MODULE
304
HYSTERESIS MODULE
306
DURATION MODULE
308
PERSISTENCE MODULE
310
ROUTE MODULE
102
DATA ANALYSIS COMPONENT

FIG. 3

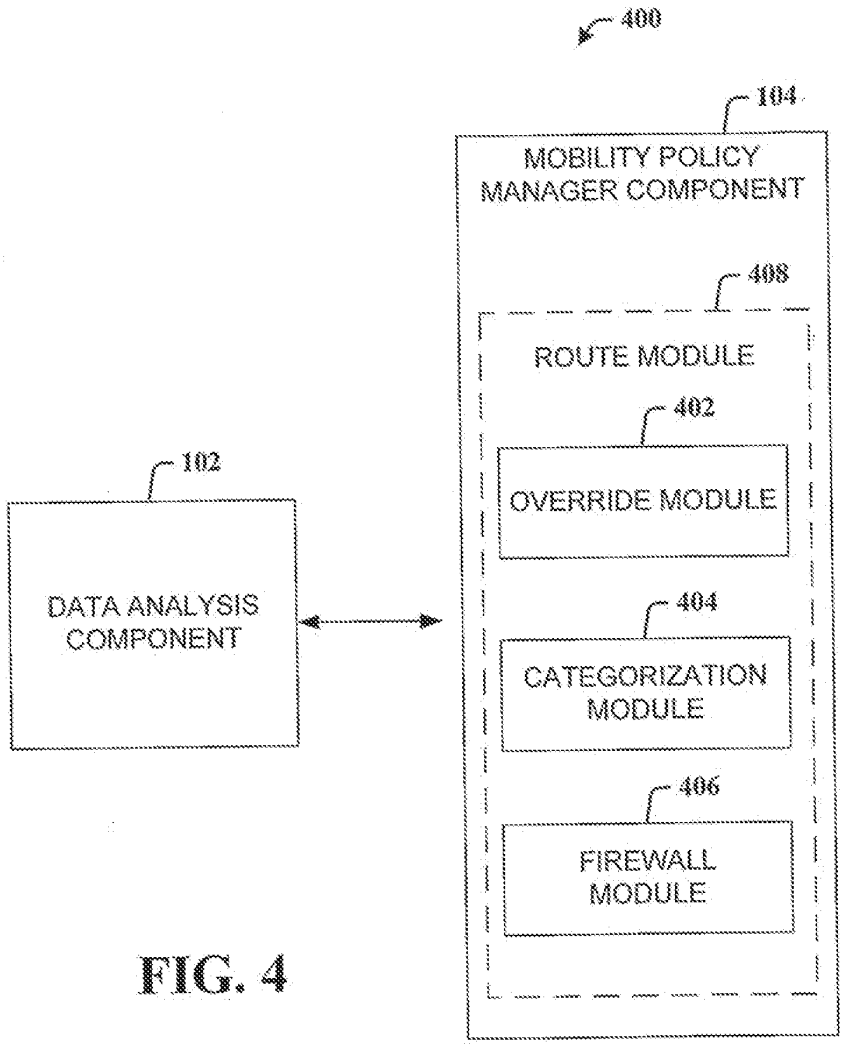

400
104
MOBILITY POLICY MANAGER COMPONENT
408
ROUTE MODULE
402
OVERRIDE MODULE
404
CATEGORIZATION MODULE
406
FIREWALL MODULE
102
DATA ANALYSIS COMPONENT

FIG. 4

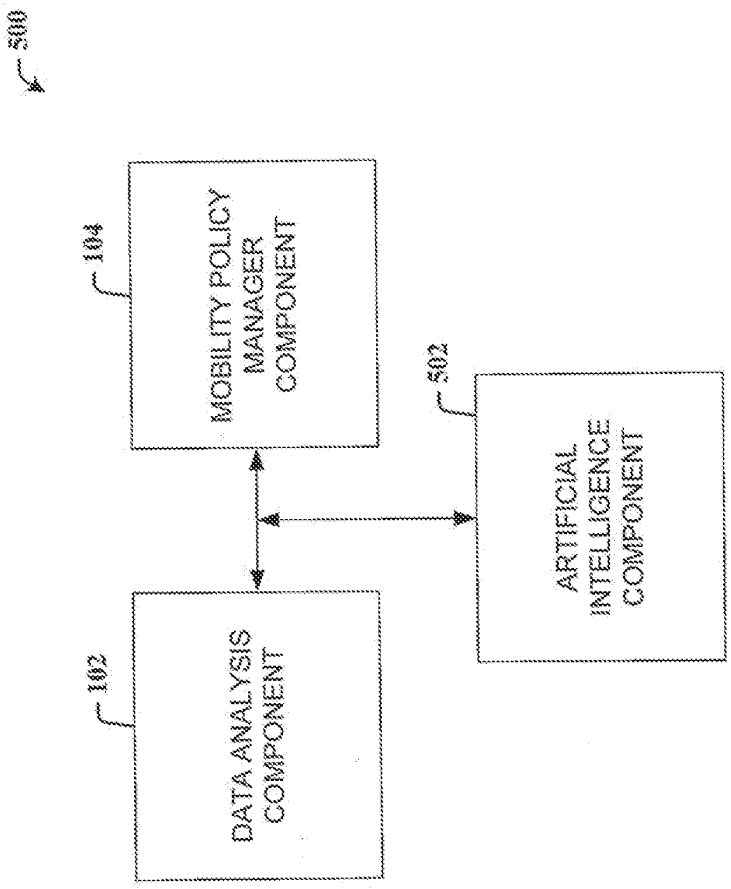

500
104
MOBILITY POLICY MANAGER COMPONENT
502
ARTIFICIAL INTELLIGENCE COMPONENT
102
DATA ANALYSIS COMPONENT

FIG. 5

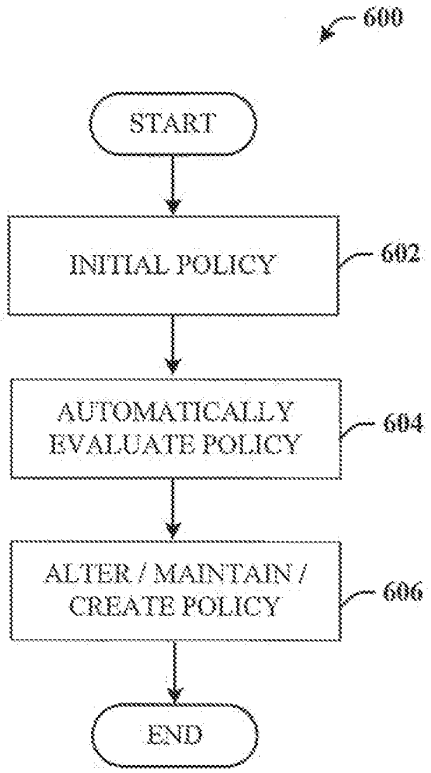

600
START
INITIAL POLICY
602
AUTOMATICALLY EVALUATE POLICY
604
ALTER / MAINTAIN / CREATE POLICY
606
END

FIG. 6

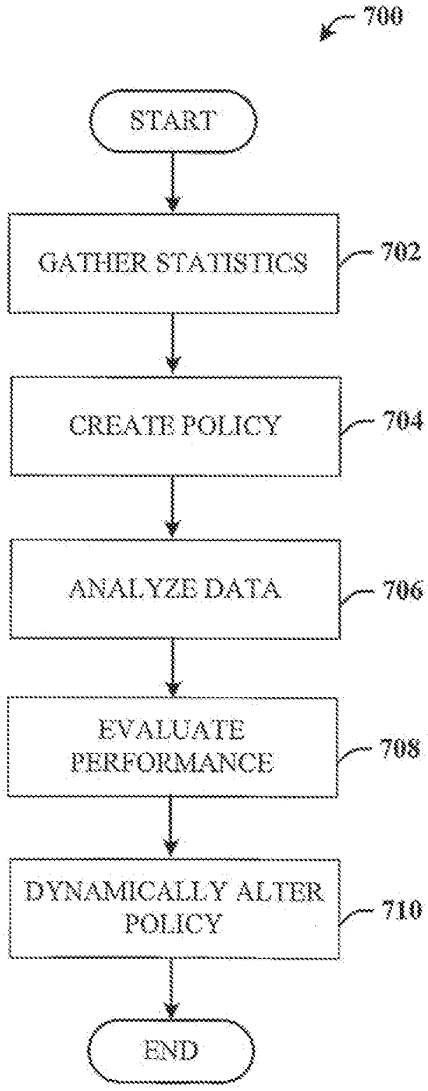

700
START
GATHER STATISTICS
702
CREATE POLICY
704
ANALYZE DATA
706
EVALUATE PERFORMANCE
708
DYNAMICALLY ALTER POLICY
710
END

FIG. 7

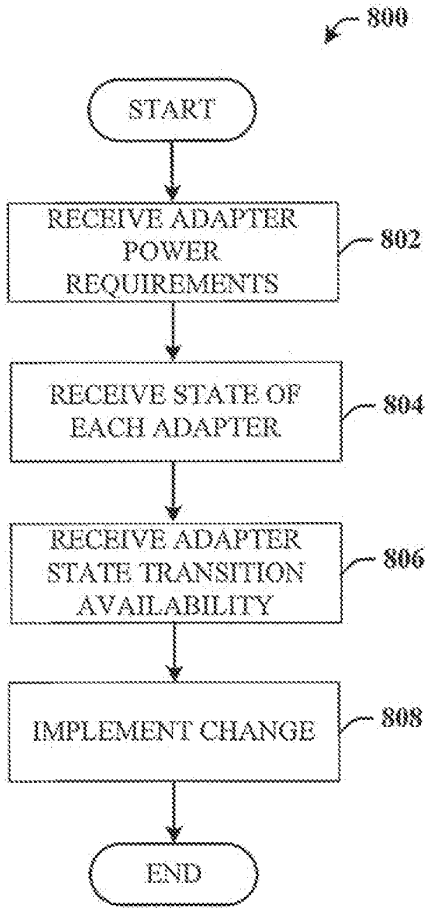

800
START
RECEIVE ADAPTER POWER REQUIREMENTS
802
RECEIVE STATE OF EACH ADAPTER
804
RECEIVE ADAPTER STATE TRANSITION AVAILABILITY
806
IMPLEMENT CHANGE
808
END

FIG. 8

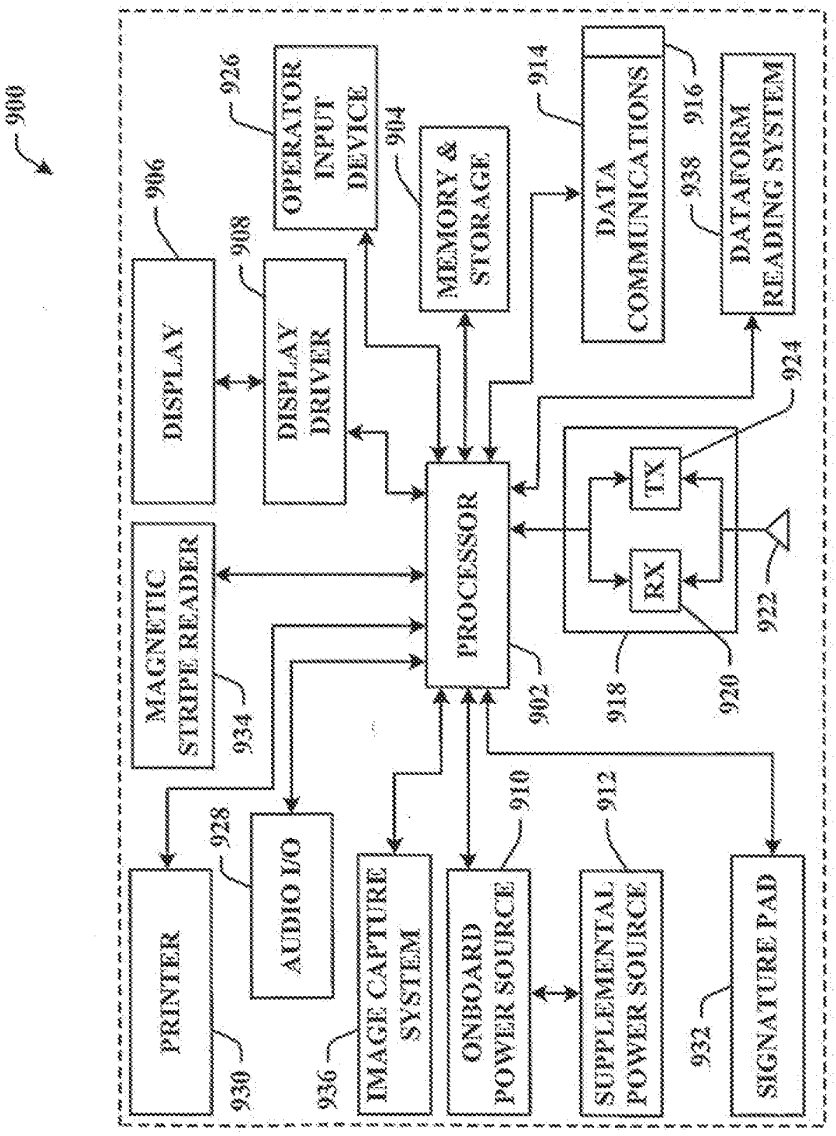

900
PRINTER
930
928
AUDIO I/O
936
IMAGE CAPTURE SYSTEM
ONBOARD POWER SOURCE
910
SUPPLEMENTAL POWER SOURCE
912
932
SIGNATURE PAD
MAGNETIC STRIPE READER
934
906
DISPLAY
908
DISPLAY DRIVER
926
OPERATOR INPUT DEVICE
904
MEMORY & STORAGE
PROCESSOR
902
914
DATA COMMUNICATIONS
938
916
DATAFORM READING SYSTEM
924
TX
RX
918
920
922

FIG. 9

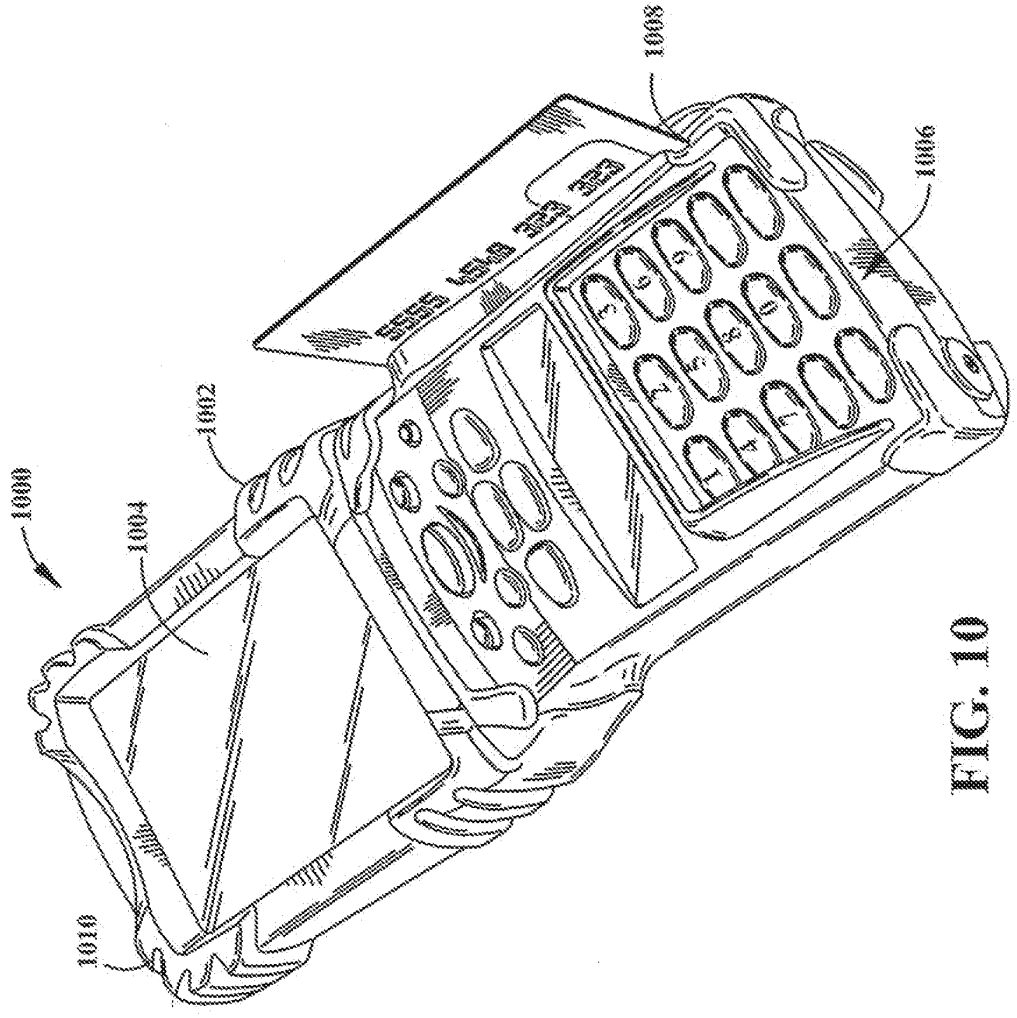


FIG. 10

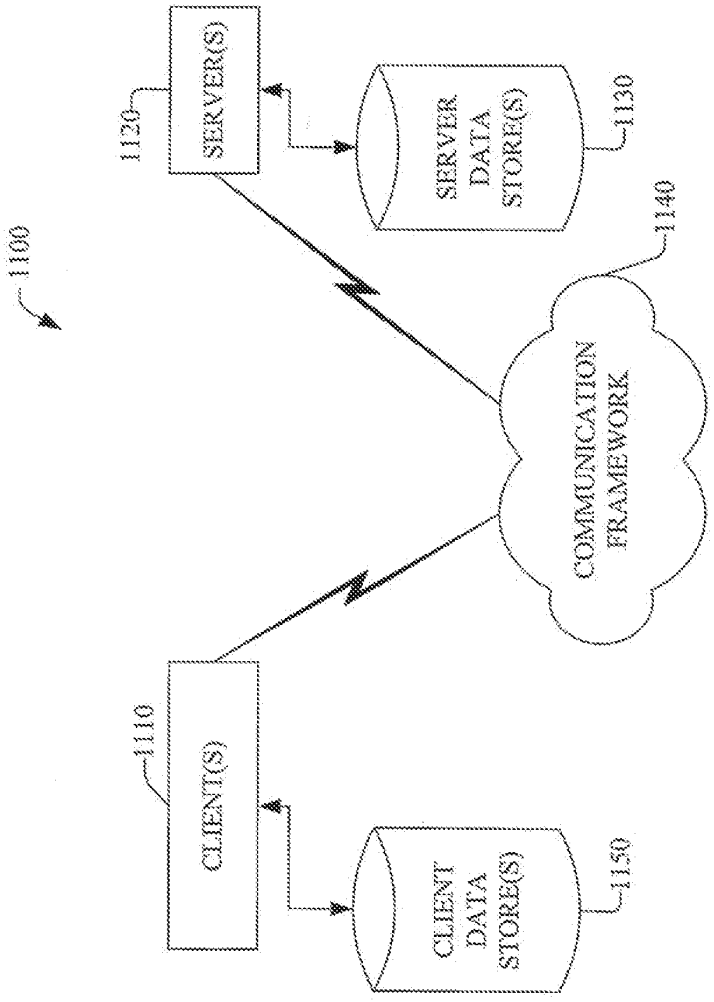

1100
1120
SERVER(S)
SERVER DATA STORE(S)
1130
1140
COMMUNICATION FRAMEWORK
1110
CLIENT(S)
CLIENT DATA STORE(S)
1150

FIG. 11

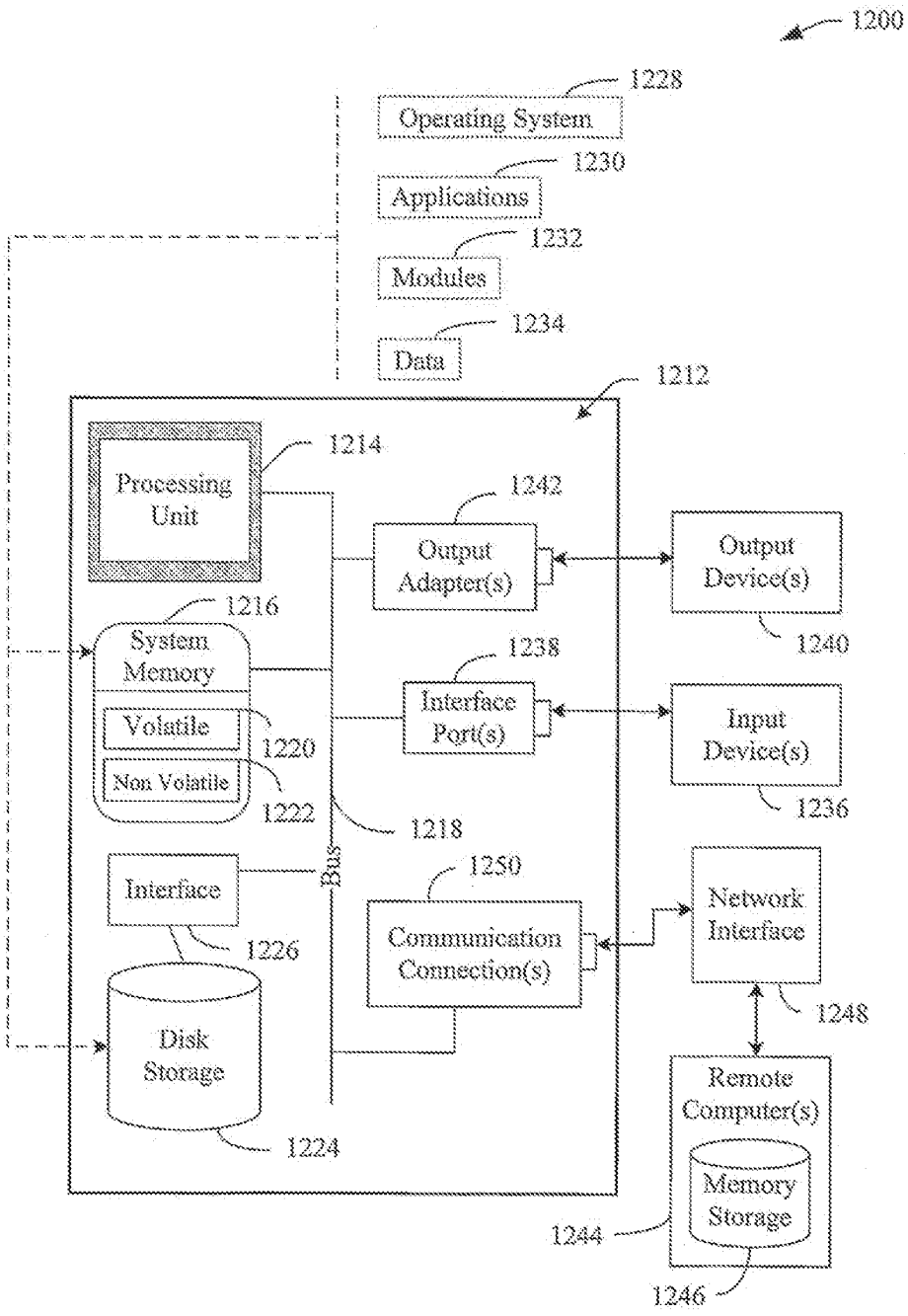

1200
1228
Operating System
1230
Applications
1232
Modules
1234
Data
1212
1214
Processing Unit
1242
Output Adapter(s)
Output Device(s)
1240
1216
System Memory
Volatile
1220
Non Volatile
1222
1238
Interface Port(s)
Input Device(s)
1236
1218
Bus
Interface
1226
1250
Communication Connection(s)
Network Interface
1248
Disk Storage
1224
Remote Computer(s)
Memory Storage
1244
1246

FIG. 12

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 739 518 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
03.01.2007 Bulletin 2007/01

(51) Int Cl.:
G06F 1/00 *(2006.01)*

(21) Application number: **05254068.9**

(22) Date of filing: **29.06.2005**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR LV MK YU

(71) Applicant: **Research In Motion Limited**
Waterloo, Ontario N2L 3W8 (CA)

(72) Inventors:
- **Adams, Neil P.**
  Waterloo
  Ontario N2K 4E4 (CA)
- **Little, Herbert A.**
  Waterloo
  Ontario N2T 2V8 (CA)

(74) Representative: **Roberts, Gwilym Vaughan et al**
Kilburn & Strode
20 Red Lion Street
London WC1R 4JP (GB)

(54) System and method for privilege management and revocation

(57) The present disclosure relates generally to the management of privileges associated with certain applications that are accessible by users of electronic equipment, such as, for example, networked computers, mobile wireless communications devices, and the like. In particular, the disclosure is directed to systems and methods for managing privileges associated with particular applications and for revoking these privileges in a timely and robust manner. For example, the device keeps track of which applications get access to which privileges. When policies or application control changes, the system detects which privileges have been revoked for which applications. This can be accomplished by simply comparing the old set of privileges with the new set of privileges. For each revoked privilege for a given application, the system determiner if the application has ever accessed that privilege in the past. If an application has accessed a privilege that is now revoked at any time in the past, the device is reset. To ensure that privileges that may be passed between applications are not overlooked, the device is arranged to perform a reset if any revoked privilege accessible by the device is one that may be passed between applications.

EP 1 739 518 A1

Printed by Jouve, 75001 PARIS (FR)

Description

BACKGROUND

1. Field of the Invention

[0001] The present invention relates generally to the management of privileges associated with certain applications that are accessible by users of electronic equipment, such as, for example, networked computers, mobile wireless communications devices, and the like. In particular, the disclosure is directed to systems and methods for managing privileges associated with particular applications and for revoking these privileges in a timely and robust manner.

2. Related Art

[0002] It is well known that certain electronic equipment, such as, for example, networked computers, mobile wireless communications devices, and the like, include applications resident on such equipment that may have access to certain privileges that enable the applications to perform various functions. Typically, a system administrator may use IT policy and application control to set the privileges associated with various applications present on the equipment that is subject to the administrator's control. Examples of privileges may include, for example, allowing an application to use inter-process communication (IPC), enabling the opening of internal and external connections, enabling the injection of browser filters, enabling Bluetooth ™ functionality, enabling use of e-mail, enabling the use of personal information management (PIM) functionality, use of application program interface (API), etc. It is important for the system administrator be able to track which applications have access to which privileges, and to be able to revoke privileges on an as needed basis.

[0003] For example, if an application has access to a privilege, and the system administrator revokes that privilege, the application should immediately, or within a small window of time, be denied access to that privilege. In other words, the privilege should be revoked as soon as possible. Events that might trigger a revocation of privileges may include, for example, an application being loaded before the IT administrator/application control data is present on the device, an application is discovered to be a rogue application, or company policy changes, resulting in limiting the use or availability of certain applications and/or privileges associated therewith.

[0004] Regardless of the reason for privilege revocation, such revocation must be accomplished in a secure manner and in a manner that prevents possible work arounds by malicious applications or individuals. In general, according to current privilege revocation schemes, privilege checking is typically performed on the first access to a privilege. For example, applications communicate with IPC using the application registry. Once an application has a reference (e.g., pointer) to the application registry, it is difficult to take this reference away from the application. In another example, if an application has passed some if its privileges to another application using IPC, conventional systems can detect that the first application has access to IPC, but there is no way to detect that the other application has been passed the privilege (s).

[0005] Therefore, there remains a need for a system and method for effectively managing privileges associated with applications, and in particular, when privilege revocation is required, to revoke these privileges in a timely and robust manner.

SUMMARY

[0006] In view of the foregoing, we have now identified an efficient, accurate, robust and easy to imaplement system and method for managing privileges associated with particular applications and for revoking these privileges in a timely and reliable manner

[0007] According to an exemplary embodiment an application having privileges that are to be revoked may be shut down. When this application is restarted, access to revoked privileges will be denied. In order to keep track of which applications need to be shut down and reset, the system must keep track of which applications have accessed which privileges. To enable monitoring of the privileges accessed by applications, each time an application uses a privilege, the system records the application identifier and which privilege has been accessed. This may be done in any of a number of conventional methods, such as, for example, a data table listing the application identifier and corresponding accessed privilege(s) associated with the application identifier. When the system administrator, or any other authority, institutes a change in privileges, the system accesses the record of which applications have accessed which privileges. A comparison of accessed privileges with the privileges to be revoked is performed. Each application that has accessed a privilege that is to be revoked is identified by the system and then shut down. When the application is restarted, the application will not have access to any of the revoked privileges.

[0008] In another embodiment, a device reset may be performed. By resetting a device, the system is brought to a known state. According to this embodiment, a system administrator, or any other authority, specifies that a device must be reset when a new policy that revokes (or changes) privileges is instituted. A device reset may be instituted at any time the administrator or authority deems it necessary to do so. For example, a device reset may be done whenever a new policy that includes privilege revocation is instituted. However, this may result in numerous unnecessary device reset events that are potentially inconvenient to the user and may interfere with use of the device. Alternatively, the system may keep track of which devices include which applications, and use this

information to determine a less intrusive device reset schedule. For example, resetting only those devices having applications that may potentially be affected by the policy change.

[0009] In yet another advantageous embodiment, aspects of previously described embodiments are combined to provide effective privilege management and revocation. According to this exemplary embodiment, the system administrator or other authority has no actual control over when a device is reset. The administrator merely manages the privileges of the system and particular applications. The device itself is responsible for resetting when necessary. In this embodiment, the device keeps track of which applications get access to which privileges. When policies or application control changes, the system detects which privileges have been revoked for which applications. This can be accomplished by simply comparing the old set of privileges with the new set of privileges. For each revoked privilege for a given application, the system determines if the application has ever accessed that privilege in the past. As noted above, the system has been keeping track of these since the system was first started up. If an application has accessed a privilege that is now revoked at any time in the past, the device is reset. For some privileges (e.g., IPC), it still cannot be determined whether a privilege has been used as a result of being passed from another application. To overcome the potential for missing a revocation of privileges when applications pass privileges between themselves, if, for example, IPC (or any other privilege that is able to be passed between applications) is revoked from any application, regardless of if the system has detected that the application has accessed the privilege, the device must be reset. This reset will bring the device back to a known state. Device reset will only be done when necessary, thus limiting the number of resets and solving the problem associated with privileges that have been passed between applications.

[0010] Each of these embodiments is useful in a variety of privilege management environments ranging from a low-level low-priority where a modest level of privilege revocation security is needed, resulting in a less robust system, to a highly important ultra-robust environment where user inconvenience is secondary to the need for absolute certainty of privilege revocation.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] These and other embodiments along with their attendant advantages are described herein with reference to the following drawings in which like reference numerals refer to like elements, and wherein:

FIG. 1 is a block diagram showing a computer network and communication system in which electronic devices running applications having access to associated privileges are used;

FIG. 2 is a block diagram of a wireless mobile communication device as an example of an electronic device running applications having access to associated privileges;

FIG. 3 is a flow diagram illustrating a method of revoking privileges according to an exemplary embodiment;

FIG. 4 is a flow diagram illustrating a method of revoking privileges according to another exemplary embodiment; and

FIG. 5 is a flow diagram illustrating yet another method of revoking privileges according to another exemplary embodiment.

DETAILED DESCREMON OF EXEMPLARY EMBODIMENTS

[0012] FIG. 1 is a block diagram showing a computer network and communication system in which electronic devices running applications having access to associated privileges are uses. The computer network 18 includes, for example, various networked computers 28 and, optionally, a message server 26, all linked via a Local Area Network (LAN) 30. The communication system includes a Wide Area Network (WAN) 12 coupled to a computer system 14, a wireless network gateway 16 and the LAN 30 of the computer network 18. The wireless network gateway 16 is also connected to a wireless communication network 20 in which a wireless mobile communication device 22 (hereinafter "mobile device"), is configured to operate. The entire system 10 is typically managed by, among others, a system administrator or like authority 32.

[0013] The computer system 14 may be a desktop or laptop personal computer that is configured to communicate with the WAN 12 or any other suitable network, such as, for example, the Internet. Personal computers, such as the computer system 14, typically access the Internet via an Internet Service Provider (ISP), Application Service Provider (ASP),or the like.

[0014] The LAN 30 is an example of a typical working computer network environment, in which multiple computers 28 are connected in a network. The computer network 18 is typically located behind a security firewall 24. Within the LAN 30, a message server 26, operating on a computer behind the firewall 24 may act as the primary interface for the owner of the computer network 18 to exchange messages both within the LAN 30 and with other external messaging clients via the WAN 12. Known message servers include, for example, Microsoft Outlook™, Lotus Notes™, Yahoo!™ Messenger, AOL Instant Messenger, or any other client-server or peer-to-peer, or similar messaging clients with various architectures. Messages received by the message server 26 are distributed to mailboxes for user accounts addressed in

the received messages, and are then accessed by a user through a messaging client operating on a computer system 28. The foregoing is merely an exemplary description illustrating a client-server architecture, and in no way implies that such architecture is necessary, as other suitable architectures known to those skilled in the art may be used.

[0015] Although only a message server 26 is shown in the LAN 30, those skilled in the art will appreciate that a LAN may include other types of servers supporting resources that are shared between the networked computer systems 28, and that the message server 26 may also provide additional functionality, such as dynamic database storage for data such as, but not limited to, calendars, to-do lists, task lists, e-mail and documentation. The message server 26 and electronic messaging are described for illustrative purposes only. Systems and methods for managing and revoking privileges are applicable to a wide range of electronic devices, and are in no way limited to electronic devices with messaging capabilities.

[0016] The wireless gateway 16 provides an interface to a wireless network 20. through which messages may be exchanged with a mobile device 22. Such functions as addressing of the mobile device 22, encoding or otherwise transforming messages for wireless transmission, and any other interface functions are performed by the wireless gateway 16. The wireless gateway 16 may be configured to operate with more than one wireless network 20, in which case the wireless gateway 16 also determines a most likely network for locating a given mobile device 22 and possibly track mobile devices as users roam between countries or networks.

[0017] The mobile device 22 is, for example, a data communication device, a voice communication device, a dual-mode communication device such as many modem cellular telephones having both data and voice communications functionality, a multiple-mode device capable of voice, data and other types of communications, a personal digital assistant (PDA) enabled for wireless communications, or a laptop or desktop computer system with a wireless modem.

[0018] Any computer system with access to the WAN 12 may exchange messages with the mobile device 22 through the wireless network gateway 16. Alternatively, private wireless network gateways such as wireless Virtual Private Network (VPN) routers could be implemented to provide a private interface to a wireless network. A wireless VPN router implemented in the LAN 30 provides a private interface from the LAN 30 to one or more mobile devices such as 22 through the wireless network 20. A private interface to a mobile device 22 may also effectively be extended to entities outside the LAN 30 by providing a message forwarding or redirection system that operates with the message server 26. Such a message redirection system is disclosed in United States Patent No. 6,219,694, which is hereby incorporated into this application by reference. In this type of system, incoming messages received by the message server 26 and addressed to a user of a mobile device 22 are sent through the wireless network interface, either a wireless VPN router, the wireless gateway 16, or another interface, for example, to the wireless network 20 and to the user's mobile device 22. Another alternate interface to a user's mailbox on a message server 26 may be a Wireless Application Protocol (WAP) gateway. Through a WAP gateway, a list of messages in a user's mailbox on the message server 26, and possibly each message or a portion of each message, may be sent to the mobile device 22. A wireless network 20 normally delivers messages to and from communication devices such as the mobile device 22 via RF transmissions between base stations and devices. The wireless network 20 may, for example, be a data-centric wireless network, a voice-centric wireless network, or a dual-mode network that can support both voice and data communications over the same infrastructure. Recently developed networks include Code Division Multiple Access (GDMA) networks and General Packet Radio Service (GPRS) networks. So-called third-generation (3G) networks like Enhanced Data rates for Global Evolution (EDGE) and Universal Mobile Telecommunications Systems (UMTS) are currently under development. Older data-centric networks include, but are not limited to, the Mobitex™ Radio Network ("Mobitex"), and the DataTAC™ Radio Network ("DataTAC"). Voice-centric data networks such as Personal Communication System (PCS) networks, including Global System for Mobile Communications (GSM) and Time Division Multiple Access (TDMA) systems, have been available in North America and world-wide for several years.

[0019] FIG. 2 is a block diagram of an exemplary wireless mobile communication device as an example of an electronic device. However, it should be understood that the systems and methods disclosed herein may be used with many different types of devices, such as personal digital assistants (PDAs), desktop computers, or the like.

[0020] The mobile device 500 is preferably a two-way communication device having at least voice and data communication capabilities. The mobile device 500 preferably has the capability to communicate with other computer systems on the Internet. Depending on the firnctionality provided by the mobile device, the mobile device may be referred to as a data messaging device, a two-way pager, a cellular telephone with data messaging capabilities, a wireless Internet appliance, or a data communication device (with or without telephony capabilities). As mentioned above, such devices are refened to generally herein as mobile devices.

[0021] The mobile device 500 includes a transceiver 511, a microprocessor 538, a display 522, non-volatile memory 524, random access memory (RAM) 526, auxiliary input/output (I/O) devices 528, a serial port 530, a keyboard 532, a speaker 534, a microphone 536, a short-range wireless communications subsystem 540, and may also include other device sub-systems 542. The transceiver 511 preferably includes transmit and receive

antennas 516, 518, a receiver (Rx) 512, a transmitter (Tx) 514, one or more local oscillators (LOs) 513, and a digital signal processor (DSP) 520. Within the non-volatile memory 524, the mobile device 500 includes a plurality of software modules 524A-524N that can be executed by the microprocessor 538 (and/or the DSP 520), including a voice communication module 524A, a data communication module 524B, and a plurality of other operational modules 524N for carrying out a plurality of other functions.

[0022] The mobile device 500 is preferably a two-way communication device having voice and data communication capabilities. Thus, for example, the mobile device 500 may communicate over a voice network, such as any of the analog or digital cellular networks, and may also communicate over a data network. The voice and data networks are depicted in FIG. 2 by the communication tower 519. These voice and data networks may be separate communication networks using separate infrastructure, such as base stations, network controllers, etc., or they may be integrated into a single wireless network. References to the network 519 should therefore be interpreted as encompassing both a single voice and data network and separate networks.

[0023] The communication subsystem 511 is used to communicate with the network 519. The DSP 520 is used to send and receive communication signals to and from the transmitter 514 and receiver 512, and also exchange control information with the transmitter 514 and receiver 512. If the voice and data communications occur at a single frequency, or closely-spaced set of frequencies, then a single LO 513 may be used in conjunction with the transmitter 514 and receiver 512. Alternatively, if different frequencies are utilized for voice communications versus data communications or the mobile device 500 is enabled for communications on more than one network 519, then a plurality of LOs 513 can be used to generate frequencies corresponding to those used in the network 519. Although two antennas 516, 518 are depicted in FIG. 2, the mobile device 500 could be used with a single antenna structure. Information, which includes both voice and data information, is communicated to and from the communication module 511 via a link between the DSP 520 and the microprocessor 538,

[0024] The detailed design of the communication subsystem 511, such as frequency band, component selection, power level, etc., is dependent upon the communication network 519 in which the mobile device 500 is intended to operate. For example, a mobile device 500 intended to operate in a North American market may include a communication subsystem 511 designed to operate with the Mobitex or DataTAC mobile data communication networks and also designed to operate with any of a variety of voice communication networks, such as AMPS, TDMA, CDMA, PCS, etc., whereas a mobile device 500 intended for use in Europe may be configured to operate with the GPRS data communication network and the GSM voice communication network. Other types of data and voice networks, both separate and integrated, may also be utilized with the mobile device 500.

[0025] Communication network access requirements for the mobile device 500 also vary depending upon the type of network 519. For example, in the Mobitex and DataTAC data networks, mobile devices are registered on the network using a unique identification number associated with each device. In GPRS data networks, however, network access is associated with a subscriber or user of the mobile device 500. A GPRS device typically requires a subscriber identity module ("SIM"), which is required in order to operate the mobile device 500 on a GPRS network. Local or non-network communication functions (if any) may be operable, without the SIM, but the mobile device 500 is unable to carry out functions involving communications over the network 519, other than any legally required operations, such as '911' emergency calling.

[0026] After any required network registration or activation procedures have been completed, the mobile device 500 is able to send and receive communication signals, preferably including both voice and data signals, over the network 519. Signals received by the antenna 516 from the communication network 519 are routed to the receiver 512, which provides for signal amplification, frequency down conversion, filtering, channel selection, etc., and may also provide analog to digital conversion. Analog to digital conversion of the received signal allows more complex communication functions, such as digital demodulation and decoding, to be performed using the DSP 520. In a similar manner, signals to be transmitted to the network 519 are processed, including modulation and encoding, for example, by the DSP 520 and are then provided to the transmitter 514 for digital to analog conversion, frequency up conversion, filtering, amplification and transmission to the communication network 519 via the antenna 518. Although a single transceiver 511 is shown for both voice and data communications, in alternative embodiments, the mobile device 500 may include multiple distinct transceivers, such as a first transceiver for transmitting and receiving voice signals, and a second transceiver for transmitting and receiving data signals, or a first transceiver configured to operate within a first frequency band, and a second transceiver configured to operate within a second frequency band.

[0027] In addition to processing the communication signals, the DSP 520 also provides for receiver and transmitter control. For example, the gain levels applied to communication signals in the receiver 512 and transmitter 514 may be adaptively controlled through automatic gain control algorithms implemented in the DSP 520. Other transceiver control algorithms could also be implemented in the DSP 520 in order to provide more sophisticated control of the transceiver 511.

[0028] The microprocessor 538 preferably manages and controls the overall operation of the mobile device 500. Many types of microprocessors or microcontrollers could be used here, or, alternatively, a single DSP 520

could be used to carry out the functions of the microprocessor 538. Low-level communication functions, including at least data and voice communications, are performed through the DSP 520 in the transceiver 511. High-level communication applications, including the voice communication application 524A, and the data communication application 524B are stored in the non-volatile memory 524 for execution by the microprocessor 538. For example, the voice communication module 524A may provide a high-level user interface operable to transmit and receive voice calls between the mobile device 500 and a plurality of other voice devices via the network 519. Similarly, the data communication module 524B may provide a high-level user interface operable for sending and receiving data, such as e-mail messages, files, organizer information, short text messages, etc., between the mobile device 500 and a plurality of other data devices via the network 519.

[0029] The microprocessor 538 also interacts with other device subsystems, such as the display 522, RAM 526, auxiliary I/O devices 528, serial port 530, keyboard 532, speaker 534, microphone 536, a short-range communications subsystem 540 and any other device subsystems generally designated as 542. For example, the modules 524A-N are executed by the microprocessor 538 and may provide a high-level interface between a user of the mobile device and the mobile device. This interface typically includes a graphical component provided through the display 522, and an input/output component provided through the auxiliary I/O devices 528, keyboard 532, speaker 534, or microphone 536. Additionally, the microprocessor 538 is capable of running a variety of applications that may be present in the device non-volatile memory 524, including applications that have access to various privileges, as will be described in more detail herein.

[0030] Some of the subsystems shown in FIG. 2 perform communication-related functions, whereas other subsystems may provide "resident" or on-device functions. Notably, some subsystems, such as keyboard 532 and display 522 may be used for both communication-related functions, such as entering a text message for transmission over a data communication network, and device-resident functions such as a calculator or task list or other PDA type functions.

[0031] Operating system software used by the microprocessor 538 is preferably stored in a persistent store such as the non-volatile memory 524. In addition to the operating system and communication modules 524A-N, the non-volatile memory 524 may include a file system for storing data. The non-volatile memory 524 may also include data stores for owner information and owner control information. The operating system, specific device applications or modules, or parts thereof, may be temporarily loaded into a volatile store, such as RAM 526 for faster operation. Moreover, received communication signals may also be temporarily stored to RAM 526, before permanently writing them to a file system located in the non-volatile memory 524. The non-volatile memory 524 may be implemented, for example, with Flash memory, non-volatile RAM, or battery backed-up RAM.

[0032] An exemplary application module 524N that may be loaded onto the mobile device 500 is a PIM application providing PDA functionality, such as calendar events, appointments, and task items. This module 524N may also interact with the voice communication module 524A for managing phone calls, voice mails, etc., and may also interact with the data communication module 524B for managing e-mail communications and other data transmissions. Alternatively, all of the functionality of the voice communication module 524A and the data communication module 524B may be integrated into the PIM module.

[0033] The non-volatile memory 524 preferably provides a file system to facilitate storage of PIM data items on the device. The PIM application preferably includes the ability to send and receive data items, either by itself, or in conjunction with the voice and data communication modules 524A, 524B, via the wireless network 519. The PIM data items are preferably seamlessly integrated, synchronized and updated, via the wireless network 519, with a corresponding set of data items stored or associated with a host computer system, thereby creating a mirrored system for data items associated with a particular user.

[0034] The mobile device 500 is manually synchronized with a host system by placing the mobile device 500 in an interface cradle, which couples the serial port 530 of the mobile device 500 to a serial port of the host system. The serial port 530 may also be used to insert owner information and owner control information onto the mobile device 500 and to download other application modules 524N for installation on the mobile device 500. This wired download path may further be used to load an encryption key onto the mobile device 500 for use in secure communications, which is a more secure method than exchanging encryption information via the wireless network 519.

[0035] Owner information, owner control information and additional application modules 524N may be loaded onto the mobile device 500 through the network 519, through an auxiliary I/O subsystem 528, through the short-range communications subsystem 540, or through any other suitable subsystem 542, and installed by a user in the non-volatile memory 524 or RAM 526. Such flexibility in application installation increases the functionality of the mobile device 500 and may provide enhanced on-device functions, communication-related functions, or both. For example, secure communication applications may enable electronic commerce functions and other such financial transactions to be performed using the mobile device 500.

[0036] When the mobile device 500 is operating in a data communication mode, a received signal, such as a text message or a web page download, will be processed by the transceiver 511 and provided to the microproces-

sor 538, which preferably further processes the received signal for output to the display 522, or, alternatively, to an auxiliary I/O device 528. Owner information, owner control information, commands or requests related to owner information or owner control information, and software applications received by the transceiver 511 are processed as described above. A user of mobile device 500 may also compose data items, such as email messages, using the keyboard 532, which is preferably a complete alphanumeric keyboard laid out in the QWERTY style, although other styles of complete alphanumeric keyboards such as the known DVORAK style may also be used. User input to the mobile device 500 is further enhanced with the plurality of auxiliary I/O devices 528, which may include a thumbwheel input device, a touchpad, a variety of switches, a rocker input switch, etc. The composed data items input by the user are then transmitted over the commuaication network 519 via the transceiver 511.

[0037] When the mobile device 500 is operating in a voice communication mode, the overall operation of the mobile device 500 is substantially similar to the data mode, except that received signals are output to the speaker 534 and voice signals for transmission are generated by a microphone 536. In addition, the secure messaging techniques described above might not necessarily be applied to voice communications. Alternative voice or audio I/O devices, such as a voice message recording subsystem, may also be implemented on the mobile device 500. Although voice or audio signal output is accomplished through the speaker 534, the display 522 may also be used to provide an indication of the identity of a calling party, the duration of a voice call, or other voice call related information. For example, the microprocessor 538, in conjunction with the voice communication module 524A and the operating system software, may detect the caller identification information of an incoming voice call and display it on the display 522.

[0038] A short-range communications subsystem 540 is also be included in the mobile device 500. For example, the subsystem 540 may include an infrared device and associated circuits and components, or a Bluetooth or 802.11 short-range wireless communication module to provide for communication with similarly-enabled systems and devices. Thus, owner information insertion, owner control information insertion, and application loading operations as described above may be enabled on the mobile device 500 via the serial port 530 or other short-range communications subsystem 540.

[0039] FIG. 2 represents a specific example of an electronic device in which owner control systems and methods described herein may be implemented. Implementation of such systems and methods in other electronic devices having further, fewer, or different components than those shown in FIG. 2 would occur to one skilled in the art to which this application pertains and are therefore considered to be within the scope of the present application.

[0040] FIG. 3 is a flow diagram illustrating a method of revoking privileges according to an exemplary embodiment. In this example, an application having privileges that are to be revoked may be shut down. When the application is restarted, access to revoked privileges will be denied. In order to keep track of which applications need to be shut down and reset, the system must keep track of which applications have access to which privileges. To accomplish this, for example, the system monitors and detects use of privileges by applications 300. The system may record an application identifier associated with a particular application and which privilege has been accessed by the application 302. This may be accomplished in any number of conventional methods that are readily apparent to those skilled in the art. For example, a data table listing the application identifier and having pointers directed to privilege(s) accessed by the application or associated with the application identifier. The system then continues to monitor the system for any change, for example, a change in IT policy, that might result in a revocation of privileges 304. So long as no such change is detected in decision block 304, the system continues to monitor and keep track of applications and associated privileges.

[0041] Upon a detection of a change that would result in revocation of privileges being accessed by applications in the system by the decision block 304, such as, for example, a change in IT policy, a comparison of the recorded data relating to the accessed privileges with a listing of the new privileges is performed 306. For example, the application identifiers and associated accessed privileges recorded in step 302 are compared to the new privilege list, or to a list of revoked privileges 306. The changes discussed herein are typically instituted and administered by a system administrator or other authority who has responsibility for operation and management of the system.

[0042] As a result of the comparison 306, each application that has accessed a privilege(s) to be revoked is identified 308 by the system. Upon identification of these applications, the system implements a shut down of these identified applications 310. When these applications are restarted 312, the applications will not have access to any of the revoked privileges. The system will continue to monitor and detect the accessing of privileges and associated applications as described above.

[0043] In another embodiment, as illustrated in the flow diagram of FIG. 4, a device reset may be performed. According to this example, all devices in the system are monitored 400. A system administrator or other authority specifies, for example, that whenever there is a change in system policy that requires revocation or changes in privileges 402, all devices in the system must be reset 404. Resetting the devices brings the system to a known state, i.e., a state in which the system knows which applications have access to which privileges throughout the system. Upon restarting these applications 406 after the device reset 404, the applications will no longer have

access to any of the revoked privileges. According to this example, a device reset may be instituted any time the system administrator or authority deems it necessary to do so. For example, a device reset may be done whenever a new policy that includes privilege revocation is instituted. As a result, this solution may invoke numerous (potentially unnecessary) device reset events that may be intrusive and inconvenient for the users. However, this embodiment provides very robust and timely privilege revocation, and is thus suitable to highly secure systems where privilege management is more important than user convenience.

[0044] Turning now to FIG. 5, another advantageous embodiment implementing features of both embodiments described above with respect to FIGS. 3 and 4 is illustrated. According to this example, aspects of the previously described embodiments are combined to provide highly effective and timely privilege management and revocation. In this example, the system administrator or authority has no actual control over when a device is reset. The administrator or authority merely manages the privileges of the system and of particular applications. The device itself is responsible for resetting as needed.

[0045] In this example, the device monitors which applications of the device have access to which privileges 600, and a log of privileges for the device is kept 602. The device monitors whether policies or application control changes are made in the system 604. If there is no change detected 604, the device continues to monitor applications and keep a log of privileges for the device 600, 602. If a change in policy or application control is detected in step 604, the system determines which privileges have been revoked for which applications by comparing the old set of privileges in the log with the new set of privileges received from the system administrator 606. The device then determines if any revoked privileges are present on the device 608. If revoked privileges are detected, e.g., if an application has accessed a privilege at any time in the past that has now been revoked, the device will reset 610. As described above, resetting the device brings the system to a known state in which all applications and privileges are known. After the device is reset 610, it is restarted 612. Upon restart 612, the device applications will have access to the correct privileges. Advantageously, if no revoked privileges are detected in step 608, the device performs another check to ensure that no privileges that have been passed between application have been missed. As explained above, for some privileges, e.g., IPC, it cannot be determined whether a privilege has been used as a result of being passed from another application. To overcome the potential for missing revocation of privileges when applications pass privileges between themselves, the system checks for privileges that are able to be passed between applications (e.g., IPC) 614. If a privilege that is able to be passed between applications is revoked from any application, regardless of if the system has detected that the application has accessed the privilege 614, the device must be reset 610 to bring the system to a known state. After resetting, the device is restarted 612, and will now have the only have access to the correct privileges. In this manner, device reset will only be performed when necessary, thus limiting the number of resets and solving the problem associated with privileges that go undetected due to their ability to be passed between applications.

[0046] While this disclosure describes specific exemplary embodiments, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, the exemplary embodiments described herein, are intended to be illustrative, not limiting. Various changes may be made without departing from the true spirit and full scope of the invention, as defined in the following claims.

Claims

1. A method for managing privileges associated with applications in an electronic system, comprising:

   monitoring devices in the electronic system to detect use of privileges by applications being run by the device;
   recording which applications have accessed which privileges;
   detecting a change in privileges of the system;
   comparing recorded privileges previously accessed by applications with a set of privilege changes;
   identifying any privileges to be revoked based on a result of said comparing step;
   when a privilege to be revoked is identified, shutting down an application that has previously accessed the revoked privilege; and
   restarting any application that was shut down for having accessed a revoked privilege.

2. A method according to claim 1, wherein said step of recording, compnses:

   storing an application identifier associated with an application that has accessed a privilege; and
   associating said identifier with the privilege accessed by the application.

3. A method according to claim 2, wherein said recording is stored in a data table listing the application identifier and privileges associated therewith.

4. A method according to claim 1, wherein said set of privilege changes comprises privileges to be revoked.

5. A method according to claim 1, wherein said set of privilege changes comprises a listing of all privileges available for use, and does not include privileges to

be revoked.

6. A method according to claim 1, wherein said devices comprise a wireless mobile electronic communication device.

7. A method according to claim 1, further comprising continuously repeating the steps.

8. A method according to claim 1, wherein said change in privileges is a result of a change in system IT policy.

9. A method for managing privileges in a system comprising electronic devices having applications resident on said devices, comprising:

monitoring electronic devices that are present in the system;
detecting a change in privileges available to said electronic devices; and
resetting said electronic devices in response to detection of a change in privileges.

10. A method according to claim 9, wherein said change in privileges is a result of a change in system IT policy.

11. A method according to claim 9, wherein said change in privileges includes an indication that privileges are to be revoked.

12. A method according to claim 9, wherein a system administrator determines when said resetting step is performed.

13. A method according to claim 9, wherein said step of resetting said electronic devices brings said system to a known state.

14. A method according to claim 13, further comprising restarting said electronic devices after said devices have been reset.

15. A method according to claim 9, wherein said electronic devices include a mobile wireless communications device.

16. A method for managing privileges in a system comprising a plurality of electronic devices, said method comprising:

monitoring, by each device of the system, which applications on each device have access to which privileges;
recording privileges accessible via applications resident on each device of the system;
detecting a change in privileges available to said devices in said system;
comparing an old set of privileges with a new set of privileges, said new set of privileges corresponding to said change in privileges;
determining if any privileges are to be revoked based on said comparison;
resetting each device that has access to a privilege that is to be revoked;
determining whether any privileges to be revoked are able to be passed between applications; and
when a privilege to be revoked is able to be passed between applications, resetting each device that has access to said privilege.

17. A method according to claim 16, wherein said change in privileges is a result of a change in system IT policy.

18. A method according to claim 16, wherein said change in privileges includes an indication that privileges are to be revoked

19. A method according to claim 16, wherein said step of resetting said electronic devices brings said system to a known state.

20. A method according to claim 19, further comprising restarting said electronic devices after said devices have been reset.

21. A method according to claim 16, wherein said electronic devices include a mobile wireless communications device.

22. A system for managing privileges in a network comprising a plurality of electronic devices, said system comprising:

a processor resident in each device of the system, for monitoring which applications on each device have access to which privileges;
a memory for recording privileges accessible via applications resident on each device of the system; and
wherein said processor is operable to:

detect a change in privileges available to said devices in said system;
compare an old set of privileges with a new set of privileges, said new set of privileges corresponding to said change in privileges;
determine if any privileges are to be revoked based on said comparison;
reset each device that has access to a privilege that is to be revoked; and
determine whether any privileges to be revoked are able to be passed between ap-

plications, and when a privilege to be revoked is able to be passed between applications, resetting each device that has access to said privilege.

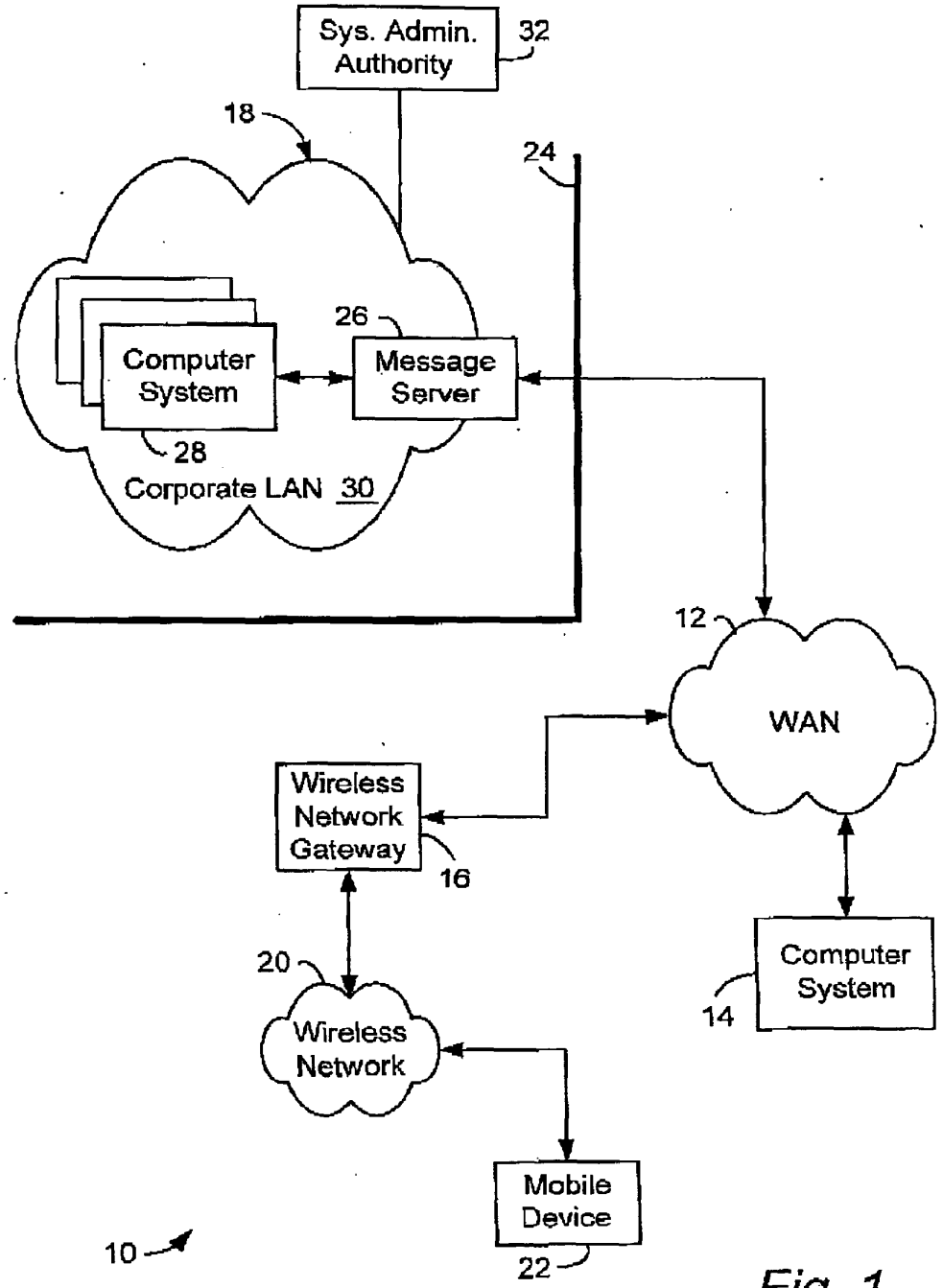

Sys. Admin. Authority
32
18
24
26
Computer System
Message Server
28
Corporate LAN 30
12
WAN
Wireless Network Gateway
16
20
Wireless Network
Computer System
14
Mobile Device
22
10

Fig. 1

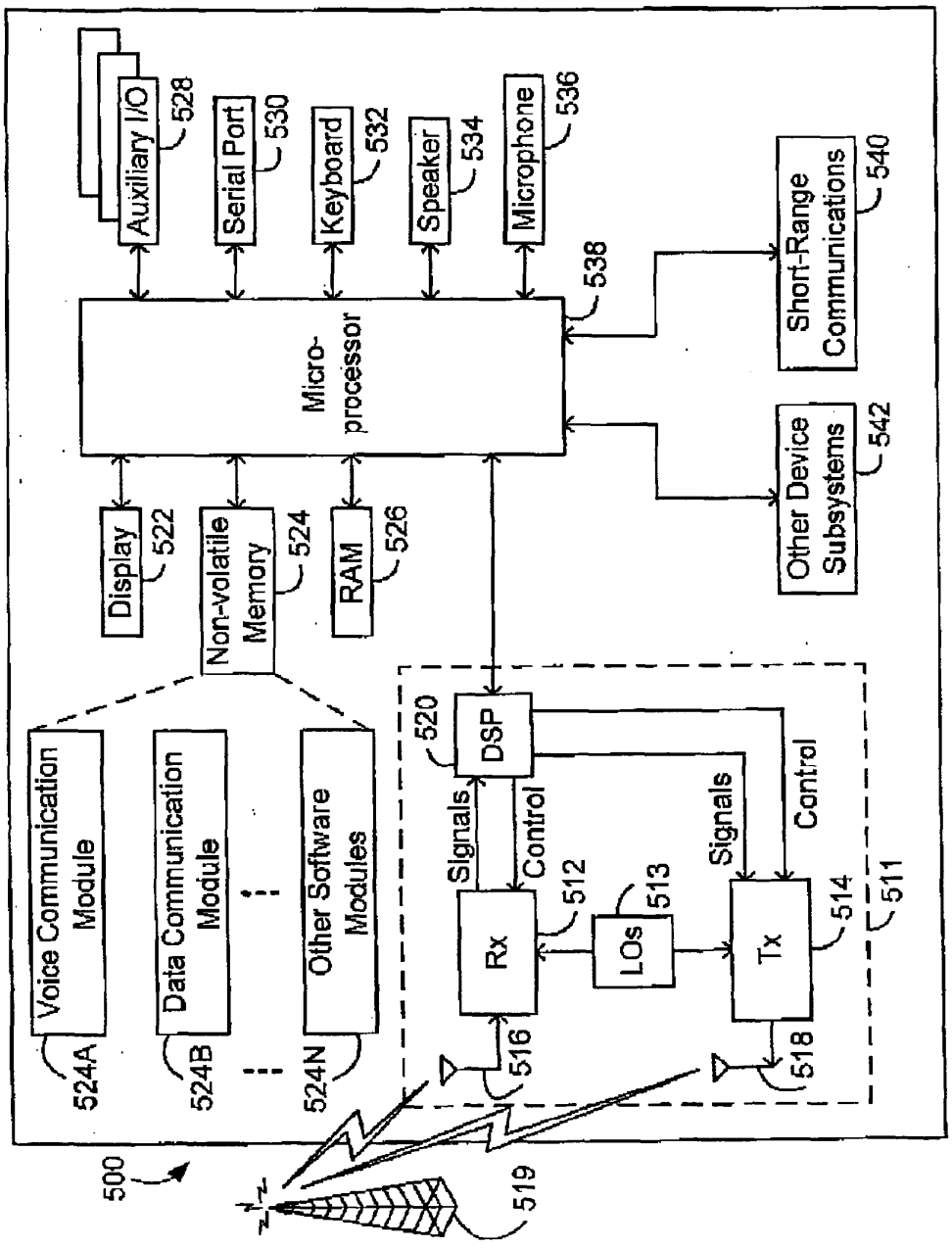

Auxiliary I/O
528
Serial Port
530
Keyboard
532
Speaker
534
Microphone
536
538
Micro-processor
Short-Range Communications
540
Other Device Subsystems
542
Display
522
Non-volatile Memory
524
RAM
526
Voice Communication Module
Data Communication Module
Other Software Modules
524A
524B
524N
520
DSP
Signals
Control
512
Rx
513
LOs
Signals
Control
514
Tx
511
516
518
500
519

Fig. 2

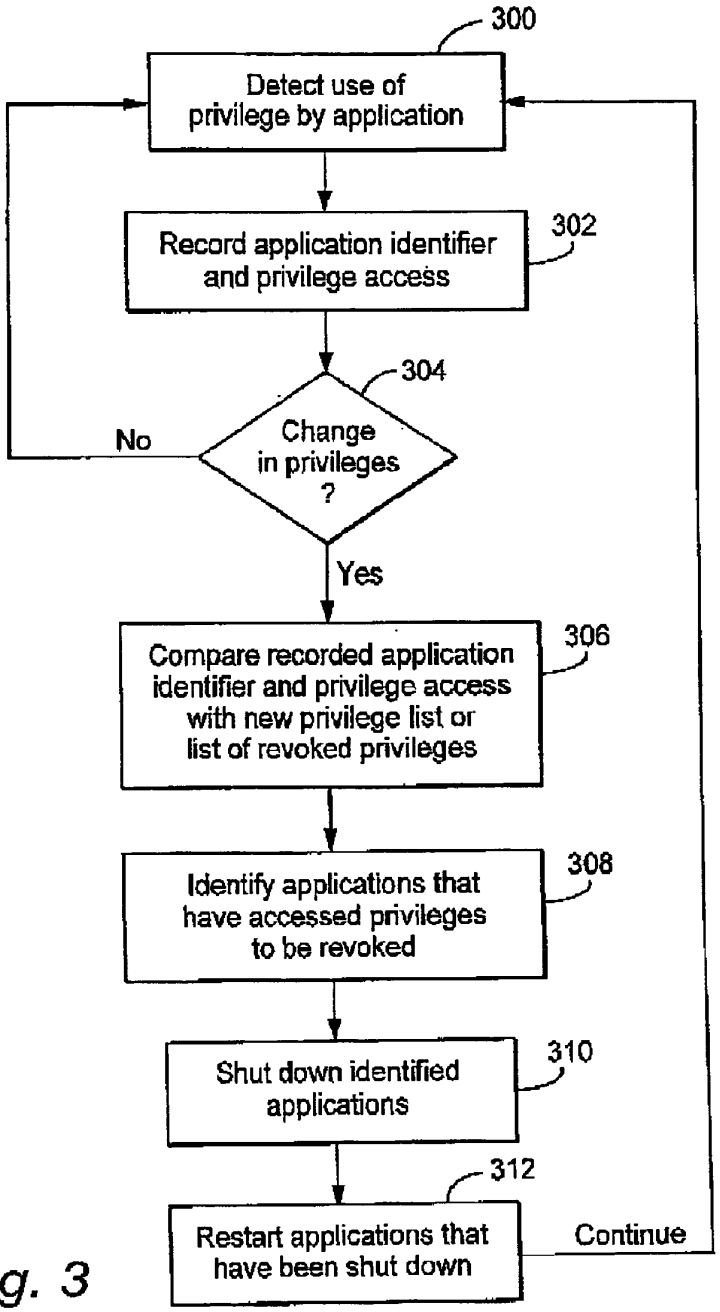

300
Detect use of privilege by application
Record application identifier and privilege access
302
304
Change in privileges ?
No
Yes
Compare recorded application identifier and privilege access with new privilege list or list of revoked privileges
306
Identify applications that have accessed privileges to be revoked
308
Shut down identified applications
310
312
Restart applications that have been shut down
Continue

Fig. 3

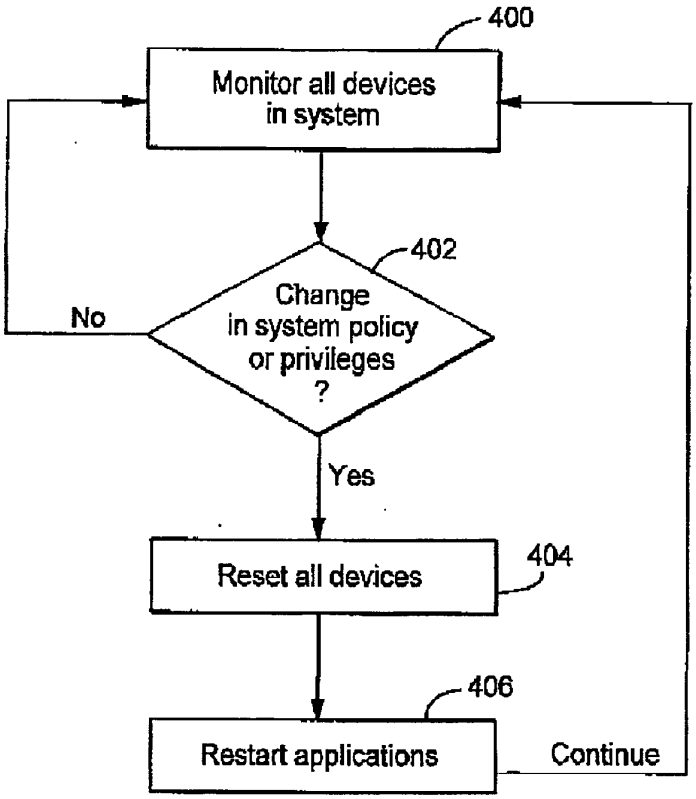

400
Monitor all devices in system
402
Change in system policy or privileges ?
No
Yes
Reset all devices
404
406
Restart applications
Continue

Fig. 4

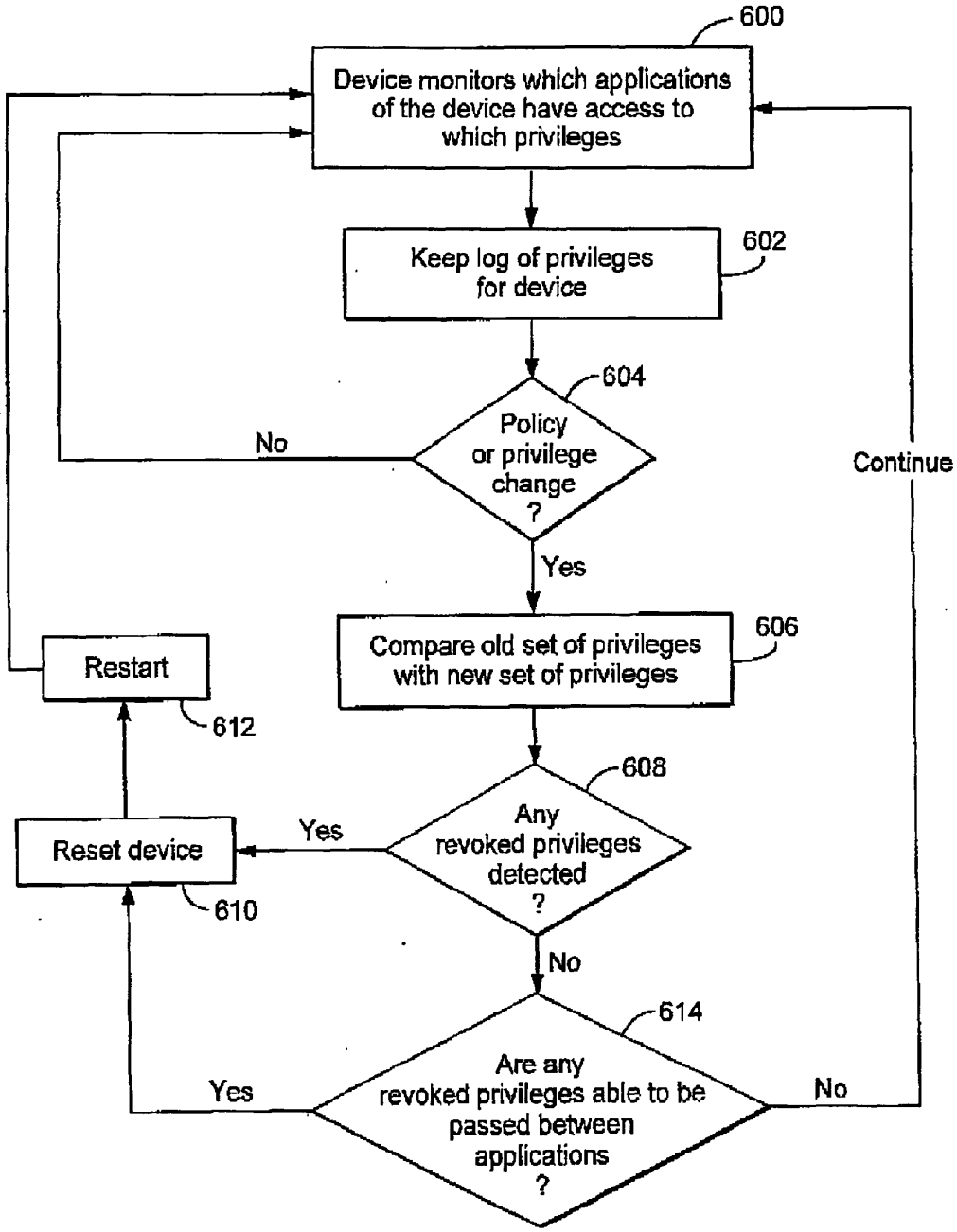

600
Device monitors which applications of the device have access to which privileges
Keep log of privileges for device
602
604
Policy or privilege change ?
No
Yes
Continue
Compare old set of privileges with new set of privileges
606
Restart
612
608
Any revoked privileges detected ?
Yes
Reset device
610
No
614
Are any revoked privileges able to be passed between applications ?
Yes
No

Fig. 5

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 05 25 4068

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	"Using software restriction policies in Windows XP and Windows .NET server to protect against unauthorized software" MICROSOFT WINDOWS XP AND WINDOES .NET TECHNICAL ARTICLE, January 2002 (2002-01), pages 1-52, XP002959638 * abstract * * page 1 * * page 4 * * pages 39-40, chapter "gpupdate.exe" *	9-15	G06F1/00
X	HANS HEDBOM, STEFAN LINDSKOG, STEFAN AXELSSON, ERLAND JONSSON: "Analysis of the Security of Windows NT" ANALYSIS OF THE SECURITY OF WINDOWS NT, 1 March 1999 (1999-03-01), XP002364916 Chalmers University of Technology, Göteborg, SWEDEN * abstract * * page 34, paragraph 5 - page 35, paragraph 5.2 *	9-15	
A	EP 1 043 648 A (SUN MICROSYSTEMS, INC) 11 October 2000 (2000-10-11) * abstract * * page 3, paragraph 10 * * page 4, paragraph 13 - paragraph 14 * * page 5, paragraph 26 * * page 6, paragraphs 35,41 * * page 8, paragraph 49 * * page 9, paragraph 58 - paragraph 64 * * page 11, paragraph 85 - page 12, paragraph 88 *	1-22	TECHNICAL FIELDS SEARCHED (IPC) G06F H04L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	27 January 2006	Bichler, M

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 05 25 4068

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

27-01-2006

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1043648 A	11-10-2000	US 6615264 B1	02-09-2003

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 6219694 B **[0018]**

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 June 2007 (21.06.2007)

PCT

(10) International Publication Number
WO 2007/069245 A2

(51) International Patent Classification:
G06F 7/04 (2006.01)

(21) International Application Number:
PCT/IL2006/001428

(22) International Filing Date:
12 December 2006 (12.12.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/750,326 13 December 2005 (13.12.2005) US
11/376,919 15 March 2006 (15.03.2006) US

(71) Applicant *(for all designated States except US)*: **YOGGIE SECURITY SYSTEMS LTD.** [IL/IL]; 42870 Beit Halevy (IL).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **TOUBOUL, Shlomo** [IL/IL]; Main 6, 42945 Kefar Haim (IL).

(74) Agents: **SANFORD T. COLB & CO.** et al.; P.O. Box 2273, 76122 Rehovot (IL).

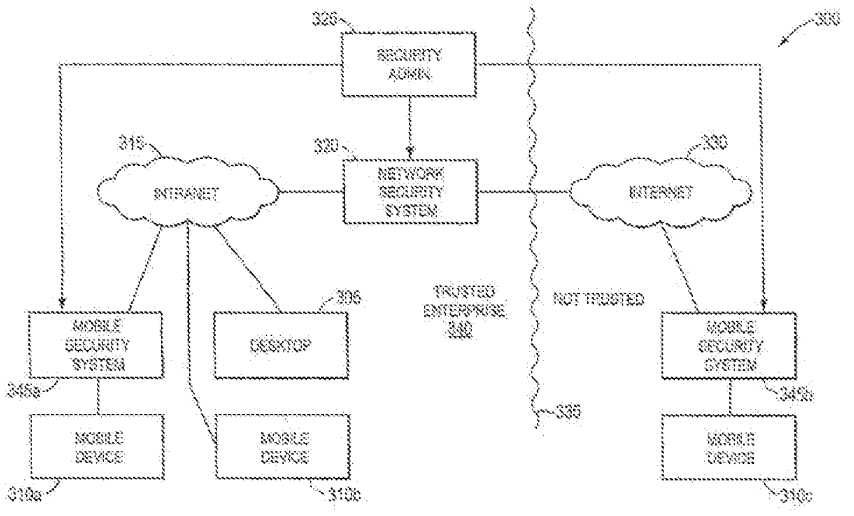
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR PROVIDING NETWORK SECURITY TO MOBILE DEVICES



(57) Abstract: A small piece of hardware connects to a mobile device and filters out attacks and malicious code. Using the piece of hardware, a mobile device can be protected by greater security and possibly by the same level of security offered by its associated corporation/enterprise. In one embodiment, a mobile security system includes a connection mechanism for connecting to a data port of a mobile device and for communicating with the mobile device; a network connection module for acting as a gateway to a network; a security policy for determining whether to forward content intended for the mobile device to the mobile device; and a security engine for executing the security policy.

WO 2007/069245 A2

SYSTEM AND METHOD FOR PROVIDING NETWORK SECURITY TO MOBILE DEVICES

PRIORITY CLAIM

This application claims benefit of and hereby incorporates by reference provisional patent application serial number 60/750,326, entitled "Personal Security Appliance," filed on December 13, 2005, by inventor Shlomo Touboul.

TECHNICAL FIELD

This invention relates generally to network security, and more particularly provides a system and method for providing network security to mobile devices.

BACKGROUND

The internet is an interconnection of millions of individual computer networks owned by governments, universities, nonprofit groups, companies and individuals. While the internet is a great source of valuable information and entertainment, the internet has also become a major source of system damaging and system fatal application code, such as "viruses," "spyware," "adware," "worms," "Trojan horses," and other malicious code.

To protect users, programmers design computer and computer-network security systems for blocking malicious code from attacking both individual and network computers. On the most part, network security systems have been relatively successful. A computer that connects to the internet from within an enterprise's network typically has two lines of defense. The first line of defense includes a network security system, which may be part of the network gateway, that includes firewalls, anti-virus, anti-spyware and content filtering. The second line of defense includes individual security software on individual machines, which is not typically as secure as the network security system and is thus more vulnerable to attacks. In combination, the first and second lines of defense together provide pretty good security protection. However, when a device connects to the internet without the intervening network security system, the device loses its first line of defense. Thus, mobile devices (e.g., laptops, desktops, PDAs such as RIM's Blackberry, cell phones, any wireless device that connects to the internet, etc.) when traveling outside the enterprise network are more vulnerable to attacks.

Figure 1 illustrates an example network system 100 of the prior art. Network system 100 includes a desktop 105 and a mobile device 110, each coupled to an enterprise's intranet 115. The intranet 115 is coupled via a network security system 120 (which may be a part of the enterprise's gateway) to the untrusted internet 130. Accordingly, the desktop 105 and mobile device 110 access the internet 130 via the network security system 120. A security administrator 125 typically manages the network security system 120 to assure that it includes the most current security protection and thus that the desktop 105 and mobile device 110 are protected from malicious code. Demarcation 135 divides the trusted enterprise 140 and the untrusted public internet 130. Because the desktop 105 and the mobile device 110 are connected to the internet

130 via the network security system 120, both have two lines of defense (namely, the network security system 120 and the security software resident on the device itself) against malicious code from the internet 130. Of course, although trusted, the intranet 115 can also be a source of malicious code.

Figure 2 illustrates an example network system 200 of the prior art, when the mobile device 110 has traveled outside the trusted enterprise 140 and reconnected to the untrusted internet 130. This could occur perhaps when the user takes mobile device 110 on travel and connects to the internet 130 at a cybercafé, at a hotel, or via any untrusted wired or wireless connection. Accordingly, as shown, the mobile device 110 is no longer protected by the first line of defense (by the network security system 120) and thus has increased its risk of receiving malicious code. Further, by physically bringing the mobile device 110 back into the trusted enterprise 140 and reconnecting from within, the mobile device 110 risks transferring any malicious code received to the intranet 115.

As the number of mobile devices and the number of attacks grow, mobile security is becoming increasingly important. The problem was emphasized in the recent Info-Security Conference in New York on December 7-8, 2005. However, no complete solutions were presented.

There is a need for personal security appliances capable of providing levels of network security as provided by enterprise network security systems.

SUMMARY

An embodiment of the present invention uses a small piece of hardware that connects to a mobile device and filters out attacks and malicious code. The piece of hardware may be referred to as a "mobile security system" or "personal security appliance." Using the mobile security system, a mobile device can be protected by greater security and possibly by the same level of security offered by its associated corporation/enterprise.

In an embodiment, a mobile security system includes a connection mechanism for connecting to a data port of a mobile device and for communicating with the mobile device; a network connection module for acting as a gateway to a network; a security policy for determining whether to forward content intended for the mobile device to the mobile device; and a security engine for executing the security policy.

The connection mechanism may include at least one of a USB connector, a PCMCIA connector, an Ethernet connector, and a BlueTooth communication module. The network connection module may include a network interface card that implements WiFi, WiMAX, GPRS, GSM, UMTS, CDMA, Generation 3, other cell phone internet connection protocols, etc. The security engine may include at least one of an antivirus engine, an antispyware engine, a firewall engine, an IPS/IDS engine, a content filtering engine, a multilayered security monitor, a bytecode monitor, and a URL monitor. The security policy may performs weighted risk analysis based on content type, content source, content category, or historical actions of the user. The remote management module may be capable of receiving security policy updates, security engine updates, and security data updates (including malicious content signatures). The mobile security system may include a distribution module capable of forwarding updates to other mobile security systems, and/or a backup module capable of storing at least a portion of the boot sector of the mobile device should the boot sector of the mobile device become compromised. The mobile security system may include a remote configuration module capable of communicating with a wizard, the wizard being in communication with an enterprise network security system, the wizard capable of substantially automatic generation of policies and data based on the policies and data on the enterprise network security system, the remote configuration module capable of

installing the policies and data generated by the wizard. The mobile security system may include a preboot memory that is not accessible during runtime, the preboot memory storing a copy of at least a portion of the operating system of the mobile security system, the mobile security system being configured to load the operating system portion every time the mobile security system is rebooted.

In another embodiment, a method comprises receiving a network connection request from a mobile device outside of a trusted network; acting as a gateway to a network on behalf of the mobile device; receiving information intended for the mobile device from the network; and determining whether to forward the information to the mobile device in accordance with a security policy.

In another embodiment, a mobile security system comprises means for acting as a gateway to a network on behalf of a mobile device outside of a trusted network; receiving information intended for the mobile device from the network; and determining whether to forward the information to the mobile device in accordance with a security policy.

In yet another embodiment, a method comprises receiving internet traffic on a mobile device via a wireless connection; redirecting the internet traffic at the kernel level to a mobile security system; scanning the internet traffic for violations of a security policy; cleaning the internet traffic of any violations of the security policy to generate cleaned internet traffic; and sending the cleaned internet traffic to the mobile device for execution.

In still another embodiment, a system comprises a wireless network interface card on a mobile device for receiving internet traffic; a kernel-level redirector on the mobile device for redirecting the internet traffic at the kernel level to a mobile security system; a security engine for scanning the internet traffic for violations of a security policy and for cleaning the internet traffic of any violations of the security policy to generate cleaned internet traffic; and a connection mechanism for receiving the redirected internet traffic from the kernel-level redirector and for sending the cleaned internet traffic to the mobile device for execution.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram of a prior art network system in a first state.

Figure 2 is a block diagram of a prior art network system in a second state.

Figure 3 is a block diagram of a network system in accordance with an embodiment of the present invention.

Figure 4 is a block diagram illustrating details of a computer system in accordance with an embodiment of the present invention.

Figure 5 is a block diagram illustrating details of the mobile security system in accordance with an embodiment of the present invention.

Figure 6 is a block diagram illustrating details of the mobile security system in accordance with a Microsoft Window's embodiment.

Figures 7 is a block diagram illustrating details of a smart policy updating system in accordance with an embodiment of the present invention.

Figure 8 is a block diagram illustrating details of network security measures relative to the OSI layers.

Figure 9 is a block diagram illustrating details of the communication technique for spreading security code to the mobile security systems.

Figures 10A-10C are block diagrams illustrating various architectures for connecting a mobile device to a mobile security system, in accordance with various embodiments of the present invention.

DETAILED DESCRIPTION

The following description is provided to enable any person skilled in the art to make and use the invention, and is provided in the context of a particular application and its requirements. Various modifications to the embodiments may be possible to those skilled in the art, and the generic principles defined herein may be applied to these and other embodiments and applications without departing from the spirit and scope of the invention. Thus, the present invention is not intended to be limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles, features and teachings disclosed herein.

An embodiment of the present invention uses a small piece of hardware that connects to a mobile device and filters out attacks and malicious code. The piece of hardware may be referred to as a "mobile security system" or "personal security appliance." Using the mobile security system, a mobile device can be protected by greater security and possibly by the same level of security offered by its associated corporation/enterprise.

Figure 3 illustrates a network system 300 in accordance with an embodiment of the present invention. Network system 300 includes a desktop 305, a first mobile device 310a, and a second mobile device 310b. The first mobile device 310a is illustrated as within the enterprise network 340 at this time and is coupled via a mobile security system 345a to the enterprise's intranet 315. The desktop 305 and second mobile device 310b are also within the enterprise network 340 but in this embodiment are coupled to the intranet 315 without an intervening mobile security system 345 such as mobile security system 345b. The intranet 315 is coupled via a network security system 320 (which may be part of the enterprise's gateway) to the untrusted internet 330. Accordingly, the first mobile device 310a, the second mobile device 310b and the desktop 305 access the untrusted internet 330 via the network security system 320. Each may also be protected by a personal security system resident thereon (not shown). A third mobile device 310c is currently outside the enterprise network 340 and is coupled via a mobile security system 345b to the untrusted internet 330. The third mobile device 310 may be in use by an employee of the trusted enterprise 340 who is currently on travel. A security administrator 325 manages the mobile security system 345a, the mobile security system 345b, and the network

security system 320 to assure that they include the most current security protection. One skilled in the art will recognize that the same security administrator need not manage the various devices. Further, the security administrator could be the user and need not be within the trusted enterprise 340.

Demarcation 335 divides the trusted enterprise 340 and the untrusted publicly accessible internet 330. Each of mobile device 310a, 310b and 310c may be referred to generically as mobile device 310, although they need not be identical. Each mobile security system 345a and 345b may be referred to generically as mobile security system 345, although they need not be identical.

As shown, although the mobile device 310c has traveled outside the trusted enterprise 340, the mobile device 310c connects to the untrusted internet 330 via the mobile security system 345b and thus retains two lines of defense (namely, the mobile security system 345b and the security software resident on the device itself). In this embodiment, the mobile security system 345 effectively acts as a mobile internet gateway on behalf of the mobile device 310c. In an embodiment, the mobile security system 345 may be a device dedicated to network security. In an embodiment, each mobile security system 345 may support multiple mobile devices 310, and possibly only registered mobile devices 310, e.g., those belonging to enterprise 340.

Each mobile security system 345 (e.g., 345a, 345b) may be a miniature server, based on commercial hardware (with Intel's Xscale as the core), Linux OS and network services, and open-source firewall, IDS/IPS and antivirus protection. The mobile security system 345 may be based on a hardened embedded Linux 2.6.

In this embodiment, because the security administrator 325 is capable of remotely communicating with the mobile security system 345b, IT can monitor and/or update the security policies/data/engines implemented on the mobile security system 345b. The security administrator 325 can centrally manage all enterprise devices, remotely or directly. Further, the security administrator 325 and mobile security systems 345 can interact to automatically translate enterprise security policies into mobile security policies and configure mobile security

systems 345 accordingly. Because the mobile security system 345 may be generated from the relevant security policies of the enterprise 340, the mobile device 310c currently traveling may have the same level of protection as the devices 305/310 within the trusted enterprise 340.

The mobile security system 345 may be designed as an add-on to existing software security or to replace all security hardware and software on a traveling mobile device. These security applications will preferably operate on different OSI layers to provide maximum security and malicious code detection, as shown in the example system illustrated in Figure 8. Operating on the lower OSI layers and doing TCP/IP packets analysis only (by screening firewall or router packets) would miss virus and/or worm behavior. Also, many modern viruses use mobile code implemented on a "higher" level than the 7th OSI layer (Application – HTTP, FTP, etc.) and therefore cannot be interpreted at the packet layer nor at the application layer. For example, applying antivirus analysis only at the session or transport layer on a malicious Java Script (that is included in an HTML page), trying to match the signature with packets and without understanding the content type (Java Script), will not detect the malicious nature of the Java Script. To offer greater protection, the mobile security system 345 may act as corporate class security appliance and engage different security applications based on the content type and the appropriate OSI layers, (or even a "higher" level if content is encapsulated in the application layer). The mobile security system 345 may be configured to perform content analysis at different OSI layers, e.g., from the packet level to the application level. It will be appreciated that performing deep inspection at the application level is critical to detect malicious content behavior and improve detection of viruses, worms, spyware, Trojan horses, etc. The following software packages may be implemented on the mobile security system 345:

- Firewall and VPN – including stateful and stateless firewalls, NAT, packet filtering and manipulation, DOS/DDOS, netfilter, isolate user mobile devices from the internet and run VPN program on the device, etc.
- Optional web accelerator and bandwidth/cache management based on Squid.
- IDS/IPS – Intrusion detection and prevention system based on Snort. Snort is an open source network intrusion prevention and detection system utilizing a rule-driven language, which combines the benefits of signature, protocol- and anomaly-based inspections.

- Antivirus and antispyware based on ClamAV; additional AV and AS engines, e.g., McAfee, Kaspersky, Pandamay, may be offered for additional subscription fees.
- Malicious Content Detection – on the fly heuristics that perform content analysis to detect malicious content before having signatures. This will be based on a rule base and updated rules and will be content dependent scanning.
- URL Categorization Filtering – based on a commercial engine, such as Surfcontrol, Smart Filters or Websense. May provide around 70 categories of URLs such as gambling, adult content, news, webmail, etc. The mobile device 345 may apply different security policies based on the URL category, e.g., higher restriction and heuristics for Gambling or Adult content web sites, etc.

FIG. 4 is a block diagram illustrating details of an example computer system 400, of which each desktop 305, mobile device 310, network security system 320, mobile security system 345, and security administrator 325 may be an instance. Computer system 400 includes a processor 405, such as an Intel Pentium® microprocessor or a Motorola Power PC® microprocessor, coupled to a communications channel 410. The computer system 400 further includes an input device 415 such as a keyboard or mouse, an output device 420 such as a cathode ray tube display, a communications device 425, a data storage device 430 such as a magnetic disk, and memory 435 such as Random-Access Memory (RAM), each coupled to the communications channel 410. The communications interface 425 may be coupled directly or via a mobile security system 345 to a network such as the internet. One skilled in the art will recognize that, although the data storage device 430 and memory 435 are illustrated as different units, the data storage device 430 and memory 435 can be parts of the same unit, distributed units, virtual memory, etc.

The data storage device 430 and/or memory 435 may store an operating system 440 such as the Microsoft Windows XP, the IBM OS/2 operating system, the MAC OS, UNIX OS, LINUX OS and/or other programs 445. It will be appreciated that a preferred embodiment may also be implemented on platforms and operating systems other than those mentioned. An embodiment may be written using JAVA, C, and/or C++ language, or other programming languages, possibly using object oriented programming methodology.

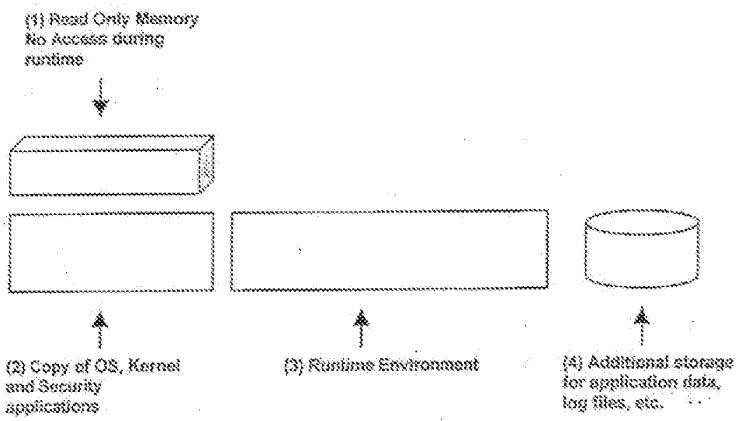
One skilled in the art will recognize that the computer system 400 may also include additional information, such as network connections, additional memory, additional processors, LANs, input/output lines for transferring information across a hardware channel, the internet or an intranet, etc. One skilled in the art will also recognize that the programs and data may be received by and stored in the system in alternative ways. For example, a computer-readable storage medium (CRSM) reader 450 such as a magnetic disk drive, hard disk drive, magneto-optical reader, CPU, etc. may be coupled to the communications bus 410 for reading a computer-readable storage medium (CRSM) 455 such as a magnetic disk, a hard disk, a magneto-optical disk, RAM, etc. Accordingly, the computer system 400 may receive programs and/or data via the CRSM reader 450. Further, it will be appreciated that the term "memory" herein is intended to cover all data storage media whether permanent or temporary.

Figure 5 is a block diagram illustrating details of the mobile security system 345 in accordance with an embodiment of the present invention. Mobile security system 345 includes adapters/ports/drivers 505, memory 510, a processor 515, a preboot flash/ROM memory module 520 storing a secure version of the mobile security system's operating system and other applications, network connection module 525, security engines 530, security policies 535, security data 540, remote management module 550, distribution module 555, and backup module 560. Although these modules are illustrated as within the mobile security system 345, one skilled in the art will recognize that many of them could be located elsewhere, e.g., on the security administrator 325 or on third-party systems in communication with the mobile security system 345. The mobile security system 345 may be in a pocket-size, handheld-size or key-chain size housing, or possibly smaller. Further, the mobile security system 345 may be incorporated within the mobile device 310.

The adapters/ports/drivers 505 include connection mechanisms (including software, e.g., drivers) for USB, Ethernet, WiFi, WiMAX, GSM, CDMA, BlueTooth, PCMCIA and/or other connection data ports on the mobile security system 345. In one embodiment, the adapters/ports/drivers 505 may be capable of connection to multiple devices 310 to provide network security to the multiple devices 310.

Memory 510 and processor 515 execute the operating system and applications on the mobile security system 345. In this example, the preboot flash 520 stores the operating system and applications. At boot time, the operating system and applications are loaded from the preboot flash 520 into memory 510 for execution. Since the operating system and applications are stored in the preboot flash 520, which cannot be accessed during runtime by the user, the operating system and applications in the preboot flash 520 are not corruptible. Should the copy of the operating system and applications in memory 510 be corrupted, e.g., by malicious code, the operating system and applications may be reloaded into the memory 510 from the preboot flash 520, e.g., upon restart. Although described as stored within the preboot flash 520, the OS and applications can be securely stored within other read-only memory devices, such as ROM, PROM, EEPROM, etc.

Memory (including memory 510 and preboot flash 520) on the mobile security system 345 may be divided into zones as follows:



Upon each "hard" restart, the boot loader (resident at area 1) of the mobile security system 345 copies the kernel and security applications (a fresh unchanged copy) from area 1 to 2. This causes a clean version of the OS and applications to be loaded into area 2 each time. That way, if a special mobile security system 345 attack is developed, the attack will be unable to infect the

system, since the OS and applications are precluded from accessing memory area 1 during runtime. Further, any attack that does reach memory 510 will be able to run only once and will disappear upon a hard restart. A triggering mechanism may be available to restart the mobile security system 345 automatically upon infection detection.

The network connection module 525 enables network connection, e.g., to the internet 330 or the intranet 315 via network communication hardware/software including WiFi, WiMAX, CDMA, GSM, GPRS, Ethernet, modem, etc. For example, if the mobile device 310 wishes to connect to the internet 330 via a WiFi connection, the adapters/ports/drivers 505 may be connected to the PCI port, USB port or PCMCIA port of the mobile device 310, and the network connection module 525 of the mobile security system 345 may include a WiFi network interface card for connecting to wireless access points. Using the network connection module 425, the mobile security system 345 may communicate with the network as a secure gateway for the mobile device 310. Other connection architectures are described in Figures 10A-10C.

The security engines 530 execute security programs based on the security policies 535 and on security data 540, both of which may be developed by IT managers. Security engines 530 may include firewalls, VPN, IPS/IDS, anti-virus, anti-spyware, malicious content filtering, multilayered security monitors, Java and bytecode monitors, etc. Each security engine 530 may have dedicated security policies 535 and security data 540 to indicate which procedures, content, URLs, system calls, etc. the engines 530 may or may not allow. The security engines 530, security policies 535 and security data 540 may be the same as, a subset of, and/or developed from the engines, policies and data on the network security system 320.

To provide a higher security level provided by antivirus and antispyware software, the security engines 530 on each mobile security system 345 may implement content analysis and risk assessment algorithms. Operating for example at OSI Layer 7 and above (mobile code encapsulated within Layer 7), these algorithms may be executed by dedicated High Risk Content Filtering (HRCF) that can be controlled by a rules engine and rule updates. The HRCF will be based on a powerful detection library that can perform deep content analysis to verify real content types. This is because many attacks are hidden within wrong mime types and/or may use

sophisticated tricks to present a text file type to a dangerous active script or ActiveX content type. The HRCF may integrate with a URL categorization security engine 530 for automatic rule adjustment based on the URL category. In one embodiment, when the risk level increases (using the described mechanism) the mobile security system 345 may automatically adjust and increase filtering to remove more active content from the traffic. For example, if greater risk is determined, every piece of mobile code, e.g., Java script, VB script, etc. may be stripped out.

Three aspects for integration with corporate policy server legacy systems include rules, LDAP and active directory, and logging and reporting as discussed below. In one embodiment, a policy import agent running on the security administrator 325 will access the rule base of Checkpoint Firewall-1 and Cisco PIX Firewalls and import them into a local copy. A rule analysis module will process the important rules and will offer out-of-the-box rules and policies for mobile security systems 345. This proposed policy will offer all mobile security systems 345 a best fit of rules that conform the firewall policy of the enterprise 340. The agent will run periodically to reflect any changes and generate updates for mobile security system 345 policies 535. The LDAP and Active Directory may be integrated with the directory service to maintain mobile security system 345 security policies 535 that respond to the enterprise's directory definitions. For example, a corporate policy for LDAP user Group "G" may automatically propagate to all mobile security systems 345 in "G" group. Mobile security system 345 local logs and audit trails may be sent in accordance to a logging and reporting policy to a central log stored at the security administrator 325. Using a web interface, IT may be able to generate reports and audit views related to all mobile device 310 users, their internet experiences, and attempts to bring infected devices back to the enterprise 340. IT will be able to forward events and log records into legacy management systems via SYSLOG and SNMP Traps.

The security engines 530 may perform weighted risk analysis. For example, the security engine 530 may analyze HTTP, FTP, SMTP, POP3, IM, P2P, etc. including any traffic arriving from the internet 330. The security engine 530 may assign a weight and rank for every object based on its type, complexity, richness in abilities, source of the object, etc. The security engine 530 may assign weight based on the source using a list of known dangerous or known safe sources. The security engine 530 may assign weight to objects based on the category of the

source, e.g., a gambling source, an adult content source, a news source, a reputable company source, a banking source, etc. The security engine 530 may calculate the weight, and based on the result determine whether to allow or disallow access to the content, the script to run, the system modification to occur, etc. The security engine 530 may "learn" user content (by analyzing for a predetermined period of time the general content that the user accesses) and accordingly may create personal content profiles. The personal content profile may be used to calibrate the weight assigned to content during runtime analysis to improve accuracy and tailor weighted risk analysis for specific user characteristics.

In some embodiments, the security engines 530, security policies 535 and security data 540 may enable bypassing the mobile security system 345. The security policy 535, set by the security administrator 325, may include a special attribute to force network connection through the mobile security system 325 when outside the trusted enterprise 340. Thus, if this attribute is set "on," when a mobile device 310 attempts to connect to the internet 330 without the mobile security system 345 and not from within the trusted enterprise 340, all data transfer connections including LAN connection, USB-net, modem, Bluetooth, WiFi, etc. may be closed. The mobile device 310 may be totally isolated and unable to connect to any network, including the internet 330.

In one embodiment, to enable this, when first connecting the mobile security system 345 to the mobile device 310 using for example the USB cable (for both power and USB connection creation), the USB plug & play device driver will be sent into the mobile device 310. The installed driver may be "Linux.inf" which allows a USB-net connection for the mobile security system 345. This connection allows the mobile security system 345 to access the internet 330 via the USB port and using the mobile device 310 network connection plus additional code ("the connection client"). In a Windows example, the connection client may be installed at the NDIS level of the mobile device 310 above all the network interface cards of every network connection as shown in Figure 6. The implementation will be as an NDIS Intermediate (IM) Driver or NDIS-Hooking Filter Driver. Both implementations may be at the kernel level, so that an end user cannot stop or remove it. When starting the mobile device 310, the connection client may attempt to connect to the security administrator 325 or the network security system 320 locally

within the trusted enterprise 340. If the node is not found (finding via VPN is considered as not found in local LAN), the connection client will assume it is working from outside the trusted enterprise 340 and expects to find the mobile security system 345 connected, e.g., via USB-net or other connection mechanism. If the mobile security system 345 is not found, the connection client may avoid any communication to any network connection. By a policy definition, this behavior can be modified to allow communication to the enterprise 340 via VPN installed in the mobile device 310. Similarly, in case of a mobile device system 345 failure, all traffic may be disabled, except for the VPN connection into the enterprise 340.

It will be appreciated that NDIS is one possible implementation of intercepting traffic at the kernel level. For example, in another embodiment, the system may hook Winsock or apply other ways that may be in future Windows versions.

In an embodiment where the mobile security system 345 supports multiple mobile devices 310, the security engines 530, security policies 535 and security data 540 may be different for each mobile device 310 (e.g., based on for example user preferences or IT decision). Alternatively, it can apply the same engines 530, policies 535 and data 540 for all connected devices 310.

The remote management module 550 enables communication with security administrator 325 (and/or other security administrators), and enables local updating of security engines 530, security policies 535, security data 540 including signatures and other applications. In one embodiment, modification to the security policies 535 and data 540 can be done by the security administrator 325 only. The remote management module 550 of the mobile security system 345 may receive updates from an update authorities device (UAD), e.g., on the security administrator 325 via a secured connection. A UAD may operate on an update server at a customer IT center located on the internet 330 to forward updates to mobile security systems 345 that possibly do not belong to an enterprise 540 in charge of managing updates. A UAD may operate on a mobile security system 345. Security engine 530 updates may modify the antivirus engine DLL, etc. OS and security application updates may be implemented only from within the enterprise 540

while connecting to the security administrator 325 and via an encrypted and authenticated connection.

The security administrator 325 can modify URL black and white lists for remote support to traveling users. In case of false positives, the security administrator 325 may allow access to certain URLs, by bypassing the proactive heuristics security but still monitoring by firewall, antivirus, IPS/IDS, etc. Additional remote device-management features may enable the security administrator 325 to perform remote diagnostics, access local logs, change configuration parameters, etc. on the mobile security system 345. The security administrator 325 may delegate tasks to a helpdesk for support.

The remote management module 550 may communicate with a wizard (e.g., wizard 745), which may be on the security administrator 325, as illustrated in Figure 7, or on another system. Details of the wizard 745 and details of the communication schemes between the remote management module 550 and the wizard 745 are described below with reference to Figure 7.

The distribution module 555 enables distribution of updates, e.g., security policy 535 updates including rule updates, security data 540 updates including signature updates, security engine 530 updates, application/OS updates, etc. by the mobile security system 345 to N other mobile security systems 345. A routing table identifying the N other mobile security systems 345 to whom to forward the updates may be provided to the distribution module 555 to enable system 345 to system 345 communication. Updates may be implemented according to policies set by the security administrator 325. When forwarding updates, the distribution module 555 acts as a UAD.

Each mobile security system 345 may obtain its routing table with security information updates, periodically, at predetermined times, upon login, etc. The routing tables may be maintained on a server, e.g., the security administrator 325 or another mobile security system 345. In one embodiment, the mobile security systems 345 may contact the server to retrieve the routing tables. Alternatively, the server may push the routing tables to the mobile security systems 345.

The distribution module 555 may enable rapid updates as shown in Figure 9. Currently, all commercial antivirus products available do not update devices faster than viruses spread. To assure that a new virus attack does not spread faster than for example signature updates, each mobile security system 345 may be an active UAD. In one embodiment, as shown in Figure 9, each mobile security system 345 is responsible for forwarding the signature updates to four other devices 345. As one skilled in the art will recognize, all devices 345 need to forward to the same number of other devices 345. Multiple devices 345 may be responsible for forwarding to the same device 345. When necessary, offline devices 345 being activated may poll the server, e.g., the security administrator 325, for routing table updates. Many other updating techniques are also possible.

The backup module 560 may constantly backup image and changes of the boot sector and system files of the mobile device 310 into the flash memory 520 or into another persistent memory device. That way, in case of major failure, including a loss of the system or boot sector of the mobile device 310, the mobile security system 345 may be identified as a CD-ROM during reboot and may launch the backup module (or separate program) to restore the boot sector and system files on the mobile device 310, thereby recovering the mobile device 310 without the need for IT support. In an embodiment where the network security system 345 supports multiple mobile devices 310, the backup module 560 may contain separate boot sector and system files for each of the mobile devices 310, if different.

Figure 7 is a block diagram illustrating details of a smart policy updating system 700 in accordance with an embodiment of the present invention. System 700 includes the security administrator 325 coupled to the network security system 320 and to the mobile security system 345. The network security system 320 includes security engines 705, including an antivirus engine 715, an IPS/IDS engine 720, a firewall engine 725, and other security engines. The network security system 320 also includes security policies and data 710, including antivirus policies and data 730, IPS/IDS policies and data 735, firewall policies and data 740, and other policies and data. Similarly, the mobile security system 345 includes an antivirus engine 755, an IPS/IDS engine 760, a firewall engine 765, and other engines. The mobile security system 345

also includes security policies and data 535/540, including antivirus security policies and data 770, IPS/IDS security policies and data 775, firewall security policies and data 780, and other security policies and data.

The security administrator 325 includes a wizard 745 for enabling substantially automatic initial and possibly dynamic setup of the security engines 530, security policies 535 and security data 540 on the mobile security system 345. In one embodiment, the wizard 745 may automatically load all security engines 705 and policies and data 710 of the network security system 320 as the security engines 530 and policies and data 535/540 on the mobile security system 345. In another embodiment, the wizard 745 may include all security engines 705 and policies and data 710 except those known to be irrelevant, e.g., those related to billing software used by accounting, those relating to web software running only on the web servers, etc. In another embodiment, the engines 530 would need to be loaded by an IT manager, and would not be loaded automatically by the wizard 745.

In one embodiment, the wizard 745 may determine whether the mobile security system 345 requires a particular security engine 530, e.g., an antivirus engine 755, IPS/IDS engine 760, firewall engine 765, etc. If so determined, then the wizard 745 would load the engine 530 onto the mobile security system 345. The wizard 745 would then determine which policies and data sets, e.g., some for antivirus engine 755, some for the IPS/IDS engine 760, some for the firewall engine 765, etc. are important to the mobile security system 345. The wizard 745 will then determine which of the antivirus policies and data 730 on the network security system 320 are relevant to the antivirus policies and data 770 on the mobile security system 345, which of the IPS/IDS policies and data 735 on the network security system 320 are relevant to the IPS/IDS policies and data 775 on the mobile security system 345, which of the firewall policies and data 740 on the network security system 320 are relevant to the firewall policies and data 780 on the mobile security system 345, and which of the other policies and data on the network security system 320 are relevant to the policies and data on the mobile security system 345. As stated above, the wizard 745 may determine that all security engines 705 or just a subset are needed on the mobile security system 345. The wizard 745 may determine that all policies and data 710 for a given engine type or just a subset should be forwarded. The wizard 745 may determine which

relevant policies and data 710 should be forwarded to the mobile security system 345 based on rules developed by an IT manager, based on item-by-item selection during the setup procedure, etc. Alternative to the wizard 745, an IT manager can setup the engines 530 and policies and data 535/540 on the mobile security system 345 without the wizard 745.

The security administrator 325 may also include an update authorities device 750. The update authorities device 750 may obtain security system updates (e.g., signature updates) and may send the updates to the network security system 320 and to the mobile security system 345. One skilled in the art will recognize that the updates to the network security system 320 and the updates to the mobile security system 345 need not be the same. Further, the update authorities device 750 may obtain the updates from security managers, security engine developers, antivirus specialists, etc. The update authorities device 750 may forward the updates to all network security systems 320 and all mobile security systems 345, or may forward routing tables to all mobile security systems 345 and the updates only to an initial set of mobile security systems 345. The initial set of mobile security systems 345 may forward the updates to the mobile security systems 345 identified in the routing tables in a P2P manner, similar to the process illustrated in Figure 9. As stated above, each mobile security system 345 operating to forward updates is itself acting as an update authorities device 750.

Other applications may be included on the mobile security system 345. For example, add-on applications for recurring revenue from existing customers may include general email, anti-spam, direct and secured email delivery, information vaults, safe skype and other instant messaging services, etc.

- Email Security and Anti-spam – implementation of mail relay on mobile security systems 345 (including the web security engine above) and a local spam quarantine (based on SendMail or similar process) may implement a complete mail security suite (SMTP and POP3) including anti-spam with real time indexing (via online web spam quarries). Users may have access to the quarantine to review spam messages, release messages, modify and custom spam rules, etc., via a web interface.
- Direct and Secured Email Delivery based on mail relay will allow the mobile security system 345 to send user email directly from one mobile security system 345 to another

mobile security system 345 without using in route mail servers. This allows corporate users to send emails that need not travel in the internet, thus leaving trace and duplicates on different unknown mail servers in route. This combined with the ability to use a secured pipe between two mobile security systems is valuable to corporations. Without such methodology, people could trace emails exchange without accessing to the enterprise's mail server, by tracking down copies in intermediate mail servers that were used to deliver the messages.

* Information Vault – Application to encrypt and store end user information on the mobile security system 345 may be available only to authorized users via a web interface and a web server implemented on every mobile security system 345 (e.g., BOA, Apache, etc.)
* Safe Skype and Other IM – implementing an instant messaging client on the mobile security system 345 can guarantee that the instant messaging system or P2P application has no access to data on the mobile device 310. Adding a chipset of AC/97 to provide a sound interface on the mobile security system 325 could allow users to talk and receive calls directly from/to the mobile security system 325.

Although not shown, a small battery may be included with the mobile security system 345. This battery may be charged by the USB connection during runtime or using the power adapter at any time. The battery may guarantee proper shutdown, e.g., when user disconnects the USB cable from the mobile security system 345. It will be signaled by the system which will launch applications and system shutdown. This will ensure a proper state of the file system and flashing open files buffers.

A multi-layered defense and detection abilities is required. This may be done by a special code that is constantly monitoring the scanning result by different systems (antivirus, IDS/IPS, firewall, antispyware, URL category, etc.) and at different levels to build a puzzle and identify an attack even if its not recognized by each of the individual subsystems. By doing this, the mobile security system 345 will maintain and in some cases even improve the security level provided within the enterprise 540.

One available benefit of the mobile security system 345 is its ability to enforce the policy of the enterprise 540 on the end user while they are traveling or working from home. Since the mobile security system 345 uses similar security engines and policy as when connected from within the enterprise 540 and since the end user cannot access the internet 330 without it (except via VPN connection into the enterprise 540), IT may be capable of enforcing its security policy beyond the boundaries of the enterprise 540. The OS may be under the entire supervision of IT, while the mobile security system 345 OS acts as an end user OS under his control. This resolves the problems of who controls what and how security and productivity face minimal compromise.

A standalone version of the mobile security system 345 may offer the same functionality, and may provide a local management interface via web browser. Attractive to home users and small offices that lack an IT department, the mobile security system 345 enables the end user to launch a browser, connect to the mobile security system 345, set the different policies (update policy, security rules, etc.) including modifying the white and black URL lists, etc. There is also an opportunity to provide end users with a service of remote management of the mobile security systems 345 by subscription.

Figures 10A, 10B and 10C illustrate three example architectures of connecting a mobile security system 345 to a mobile device 310, in accordance with various embodiments of the present invention. In Figure 10A, the mobile device 310 is coupled to the mobile security system 345 via USB connections 1015 and 1020 and is coupled to the internet 330 via a NIC card 1005. The mobile device 310 receives internet traffic from the internet 330 via its NIC card 1005. A kernel-level redirector 1010 (e.g., via NDIS, Winsock, etc.) on the mobile device 310 automatically redirects the internet traffic via the USB connections 1015 and 1020 to the mobile security system 345, which scans, cleans and returns the cleaned internet traffic to the mobile device 310 via the USB connections 1015 and 1020. In Figure 10B, the mobile device 310 is coupled to the mobile security system 345 via USB connections 1025 and 1030. The mobile security system 345 includes a NIC card 1035 for receiving internet traffic from the internet 330. The mobile security system 345 scans, cleans and forwards the internet traffic via the USB connections 1025 and 1030 to the mobile device 310. In Figure 10C, the mobile device 310 is coupled to the mobile security system 345 via NIC cards 1040 and 1045. The mobile security

system 345 receives internet traffic from the internet 330 via its NIC card 1045. The mobile security system 345 scans, cleans and forwards the internet traffic wirelessly via the NIC cards 1040 and 1045 to the mobile device 310. Other connection architectures are also possible.

The foregoing description of the preferred embodiments of the present invention is by way of example only, and other variations and modifications of the above-described embodiments and methods are possible in light of the foregoing teaching. Although the network sites are being described as separate and distinct sites, one skilled in the art will recognize that these sites may be a part of an integral site, may each include portions of multiple sites, or may include combinations of single and multiple sites. The various embodiments set forth herein may be implemented utilizing hardware, software, or any desired combination thereof. For that matter, any type of logic may be utilized which is capable of implementing the various functionality set forth herein. Components may be implemented using a programmed general purpose digital computer, using application specific integrated circuits, or using a network of interconnected conventional components and circuits. Connections may be wired, wireless, modem, etc. The embodiments described herein are not intended to be exhaustive or limiting. The present invention is limited only by the following claims.

CLAIMS

1. A mobile security system, comprising:

a connection mechanism for connecting to a data port of a mobile device and for communicating with the mobile device;

a network connection module for acting as a gateway to a network;

a security policy for determining whether to forward content intended for the mobile device to the mobile device; and

a security engine for executing the security policy.

2. The mobile security system of claim 1, wherein the connection mechanism includes at least one of a USB connector, a PCMCIA connector, an Ethernet connector, and a wireless communication module.

3. The mobile security system of claim 1, wherein the network connection module includes a wireless network interface card.

4. The mobile security system of claim 1, wherein the security engine includes at least one of an antivirus engine, an antispyware engine, a firewall engine, an IPS/IDS engine, a content filtering engine, a multilayered security monitor, a bytecode monitor, and a URL monitor.

5. The mobile security system of claim 1, wherein the security policy performs weighted risk analysis.

6. The mobile security system of claim 5, wherein the weighted risk analysis weighs risk based on content type.

7. The mobile security system of claim 5, wherein the weighted risk analysis weighs risk based on content source.

8. The mobile security system of claim 5, wherein the weighted risk analysis weighs risk based on content source category.

9. The mobile security system of claim 5, wherein the weighted risk analysis weighs risk based on historical actions of the user.

10. The mobile security system of claim 1, further comprising a remote management module capable of receiving security policy updates.

11. The mobile security system of claim 1, further comprising a remote management module capable of receiving security engine updates.

12. The mobile security system of claim 1, further comprising security data and a remote management module capable of receiving security data updates.

13. The mobile security system of claim 1, wherein the security data includes malicious content signatures.

14. The mobile security system of claim 1, further comprising a distribution module capable of forwarding updates to other mobile security systems.

15. The mobile security system of claim 1, further comprising a backup module capable of storing at least a portion of the boot sector of the mobile device should the boot sector of the mobile device become compromised.

16. The mobile security system of claim 1, further comprising a remote configuration module capable of communicating with a wizard, the wizard being in communication with an enterprise network security system, the wizard capable of substantially automatic generation of policies and data based on the policies and data on the enterprise network security system, the remote configuration module capable of installing the policies and data generated by the wizard.

17. The mobile security system of claim 1, further comprising a preboot memory that is not accessible during runtime, the preboot memory storing a copy of at least a portion of the operating system of the mobile security system, the mobile security system being configured to load the operating system portion every time the mobile security system is rebooted.

18. A method comprising:

receiving a network connection request from a mobile device when the mobile device is outside of a trusted network;

acting as a gateway to a network on behalf of the mobile device;

receiving information intended for the mobile device from the network; and

determining whether to forward the information to the mobile device in accordance with a security policy.

19. A system comprising:

means for acting as a gateway to a network on behalf of a mobile device when the mobile device is outside of a trusted network;

receiving information intended for the mobile device from the network; and

determining whether to forward the information to the mobile device in accordance with a security policy.

20. A method comprising:

receiving internet traffic on a mobile device via a wireless connection;

redirecting the internet traffic at the kernel level to a mobile security system;

scanning the internet traffic for violations of a security policy;

cleaning the internet traffic of any violations of the security policy to generate cleaned internet traffic; and

sending the cleaned internet traffic to the mobile device for execution.

21. A system comprising:

a wireless network interface card on a mobile device for receiving internet traffic;

a kernel-level redirector on the mobile device for redirecting the internet traffic at the kernel level to a mobile security system;

a security engine for scanning the internet traffic for violations of a security policy and for cleaning the internet traffic of any violations of the security policy to generate cleaned internet traffic; and

a connection mechanism for receiving the redirected internet traffic from the kernel-level redirector and for sending the cleaned internet traffic to the mobile device for execution.

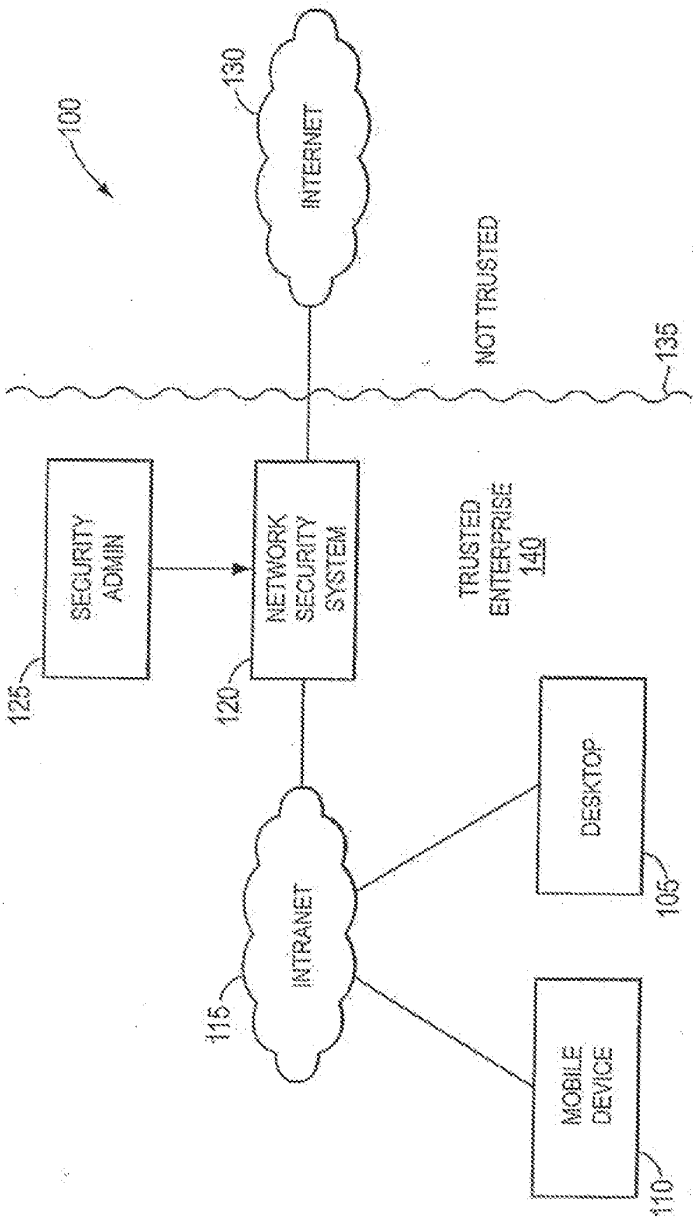


FIG. 1
(PRIOR ART)

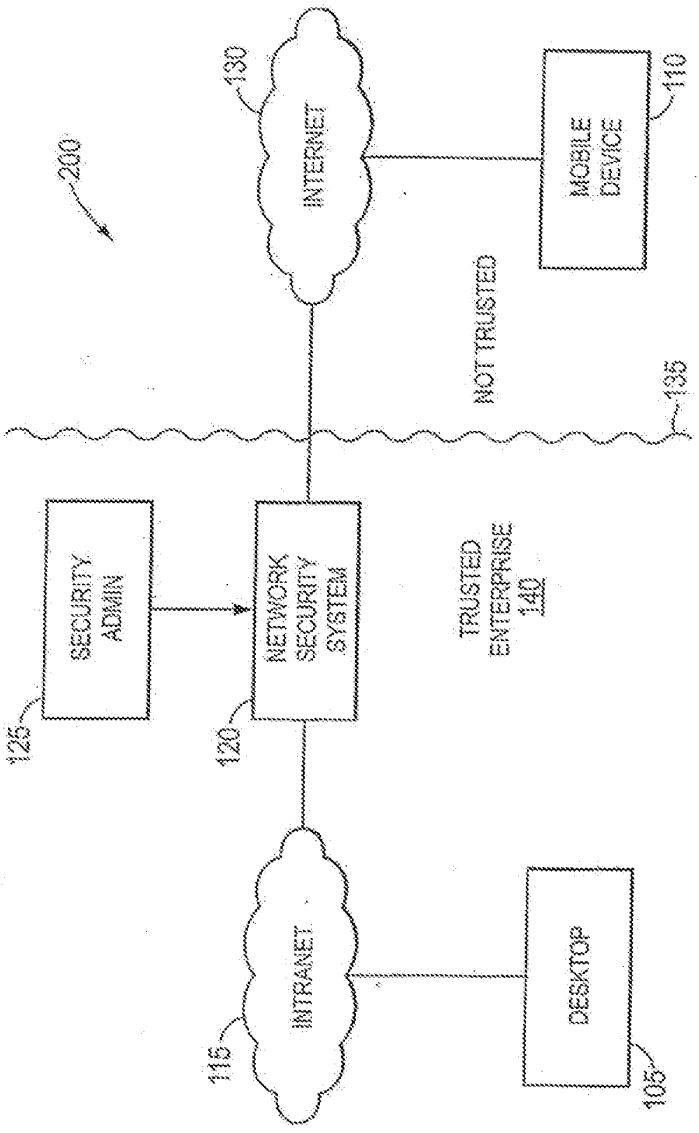


FIG. 2
(PRIOR ART)

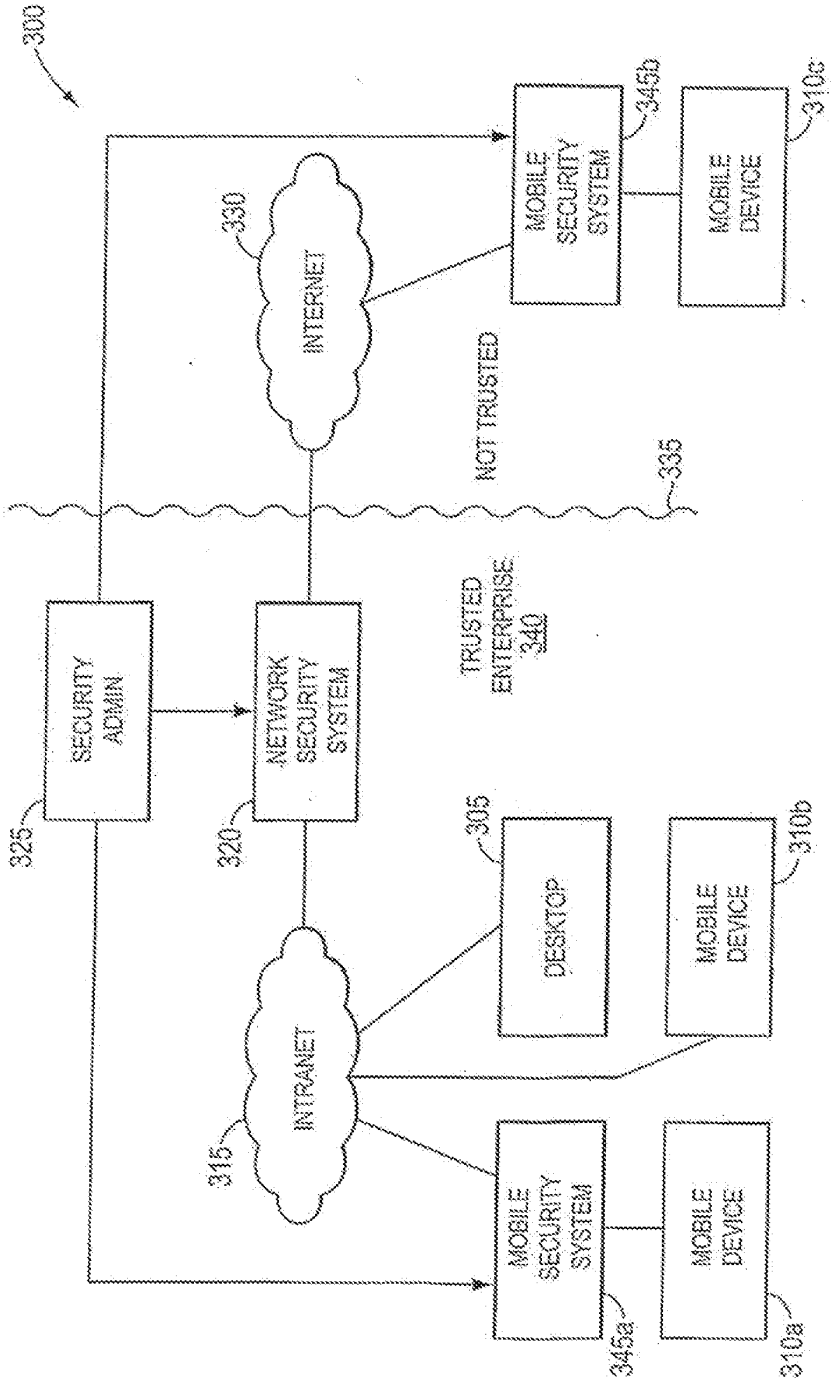

300
MOBILE SECURITY SYSTEM
345b
MOBILE DEVICE
310c
330
INTERNET
NOT TRUSTED
335
TRUSTED ENTERPRISE 340
SECURITY ADMIN
325
NETWORK SECURITY SYSTEM
320
305
DESKTOP
MOBILE DEVICE
310b
INTRANET
315
MOBILE SECURITY SYSTEM
345a
MOBILE DEVICE
310a

FIG. 3

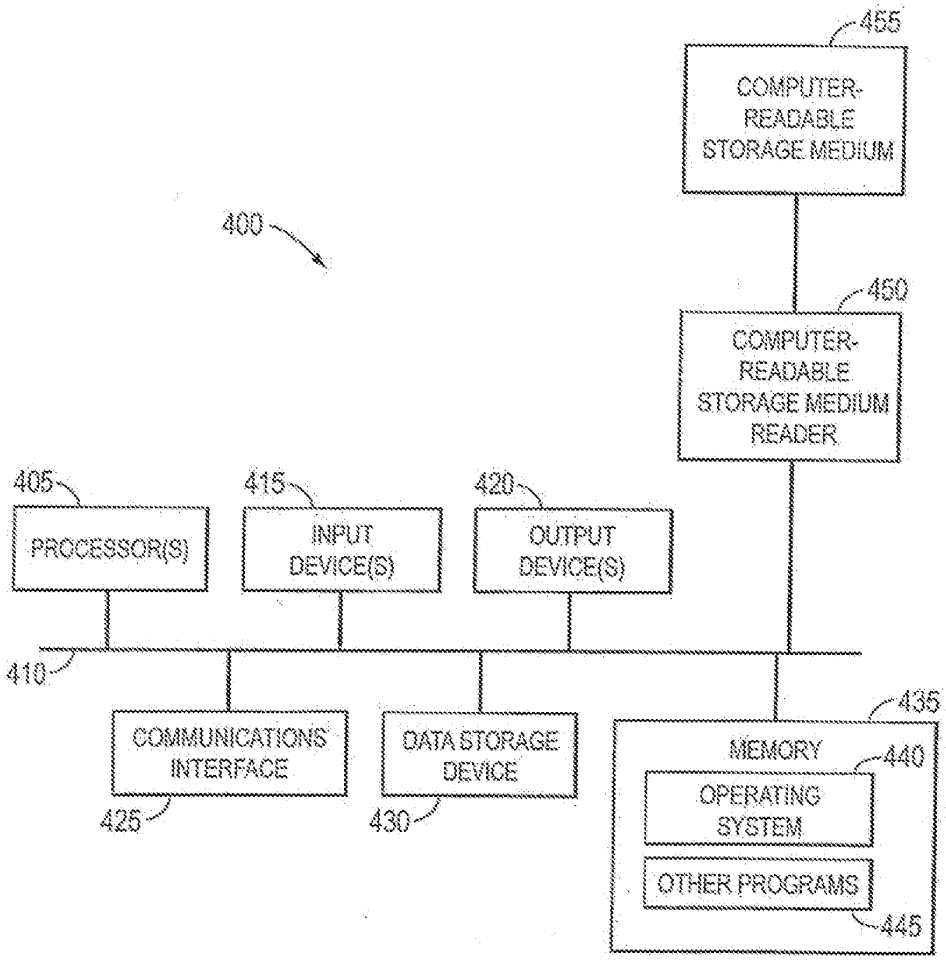

455
COMPUTER-READABLE STORAGE MEDIUM
400
450
COMPUTER-READABLE STORAGE MEDIUM READER
405
PROCESSOR(S)
415
INPUT DEVICE(S)
420
OUTPUT DEVICE(S)
410
COMMUNICATIONS INTERFACE
425
DATA STORAGE DEVICE
430
435
MEMORY
440
OPERATING SYSTEM
OTHER PROGRAMS
445

FIG. 4

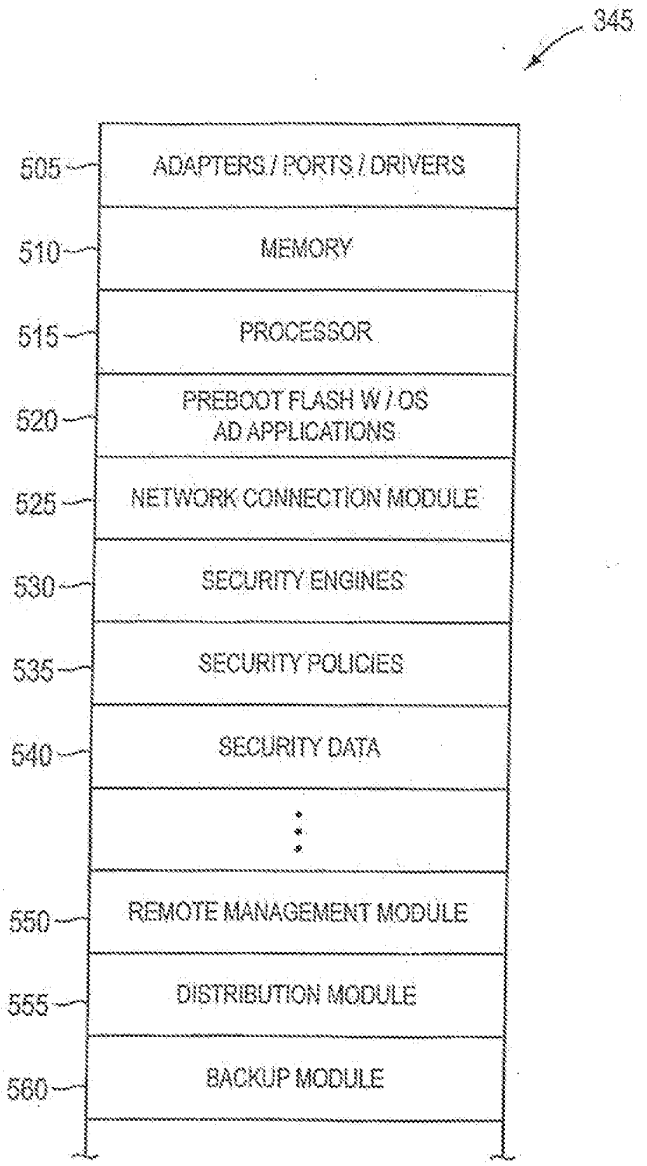

345
505
ADAPTERS / PORTS / DRIVERS
510
MEMORY
515
PROCESSOR
520
PREBOOT FLASH W / OS
AD APPLICATIONS
525
NETWORK CONNECTION MODULE
530
SECURITY ENGINES
535
SECURITY POLICIES
540
SECURITY DATA
550
REMOTE MANAGEMENT MODULE
555
DISTRIBUTION MODULE
560
BACKUP MODULE

FIG. 5

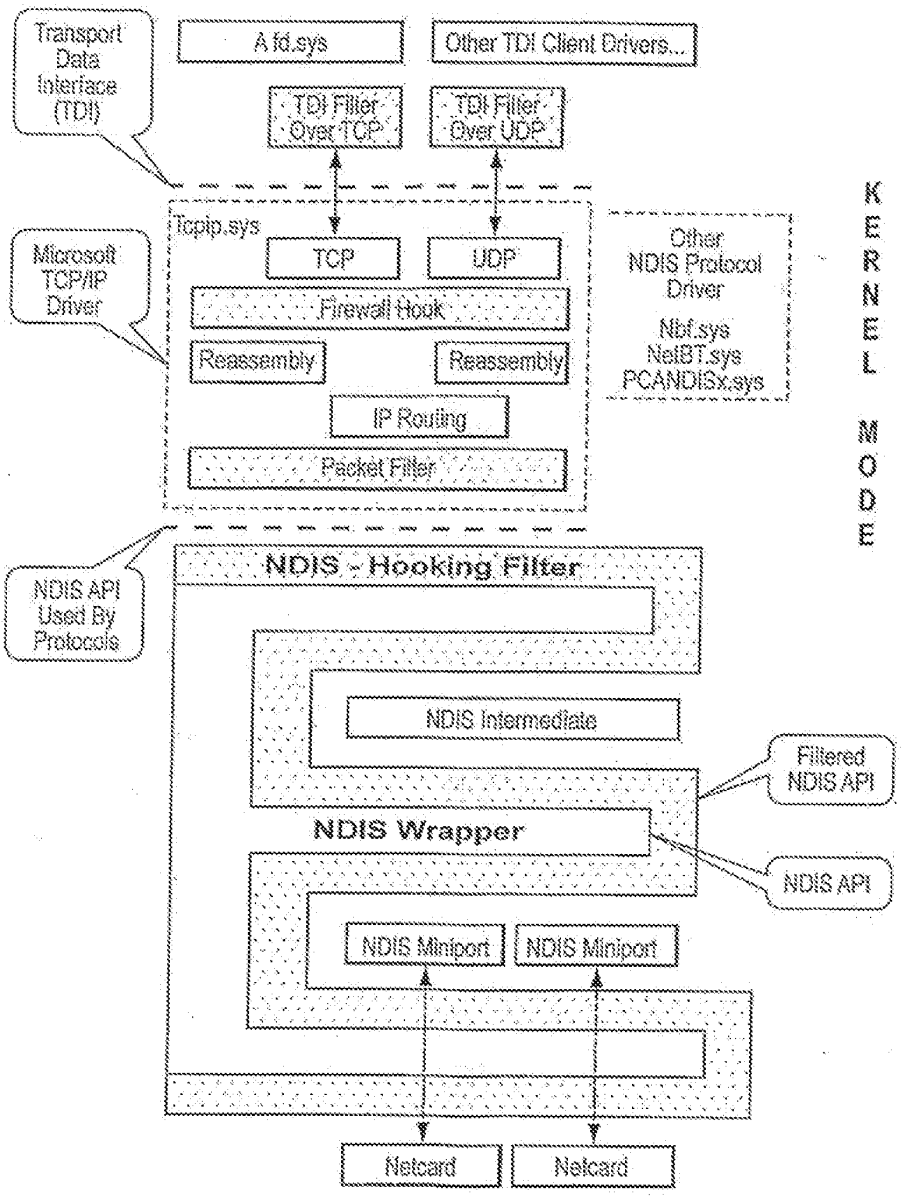

Transport Data Interface (TDI)
Afd.sys
Other TDI Client Drivers...
TDI Filter Over TCP
TDI Filter Over UDP
Tcpip.sys
TCP
UDP
Firewall Hook
Reassembly
Reassembly
IP Routing
Packet Filter
Microsoft TCP/IP Driver
Other NDIS Protocol Driver
Nbf.sys
NetBT.sys
PCANDISx.sys
KERNEL MODE
NDIS - Hooking Filter
NDIS API Used By Protocols
NDIS Intermediate
Filtered NDIS API
NDIS Wrapper
NDIS API
NDIS Miniport
NDIS Miniport
Netcard
Netcard

FIG. 6

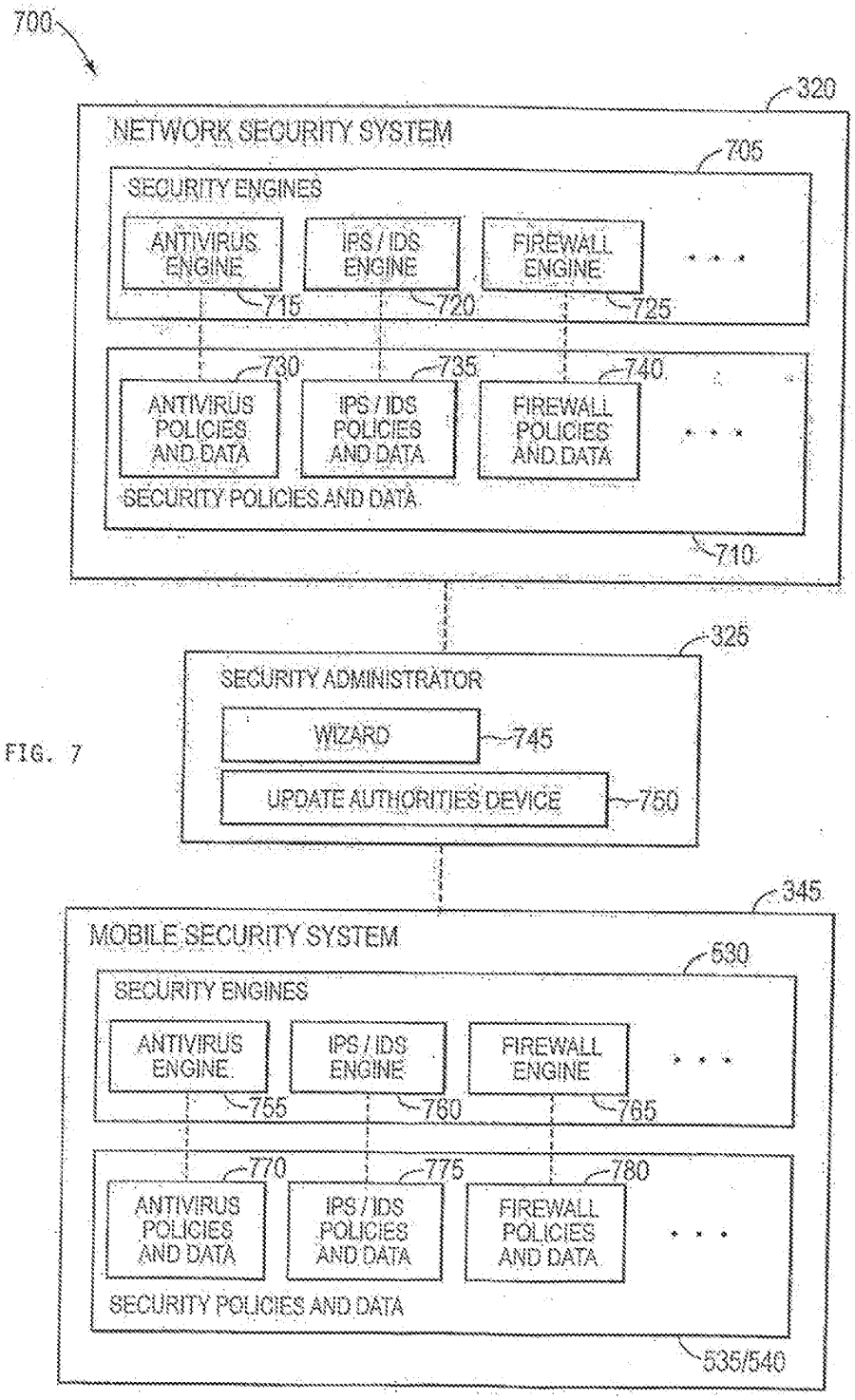

700
320
NETWORK SECURITY SYSTEM
705
SECURITY ENGINES
ANTIVIRUS ENGINE
IPS / IDS ENGINE
FIREWALL ENGINE
715
720
725
730
735
740
ANTIVIRUS POLICIES AND DATA
IPS / IDS POLICIES AND DATA
FIREWALL POLICIES AND DATA
SECURITY POLICIES AND DATA
710
325
SECURITY ADMINISTRATOR
WIZARD
745
UPDATE AUTHORITIES DEVICE
750
345
MOBILE SECURITY SYSTEM
530
SECURITY ENGINES
ANTIVIRUS ENGINE
IPS / IDS ENGINE
FIREWALL ENGINE
755
760
765
770
775
780
ANTIVIRUS POLICIES AND DATA
IPS / IDS POLICIES AND DATA
FIREWALL POLICIES AND DATA
SECURITY POLICIES AND DATA
535/540

FIG. 7

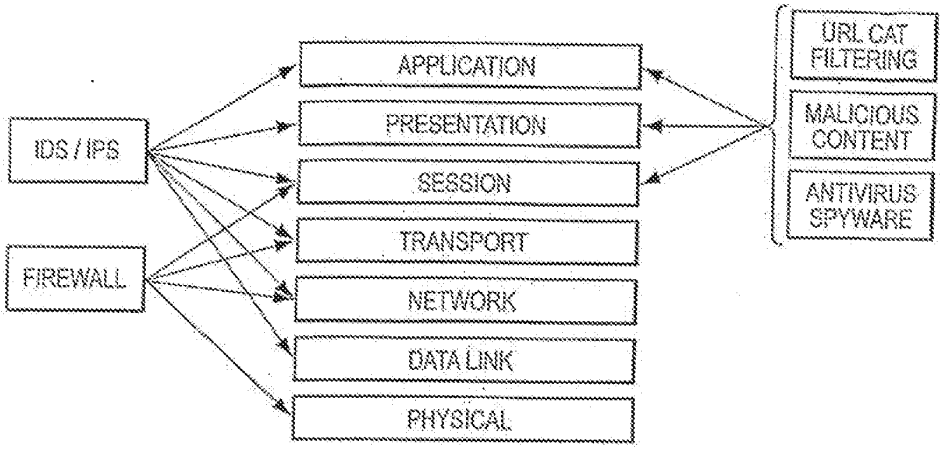

APPLICATION
PRESENTATION
SESSION
TRANSPORT
NETWORK
DATA LINK
PHYSICAL
IDS / IPS
FIREWALL
URL CAT FILTERING
MALICIOUS CONTENT
ANTIVIRUS SPYWARE

FIG. 8

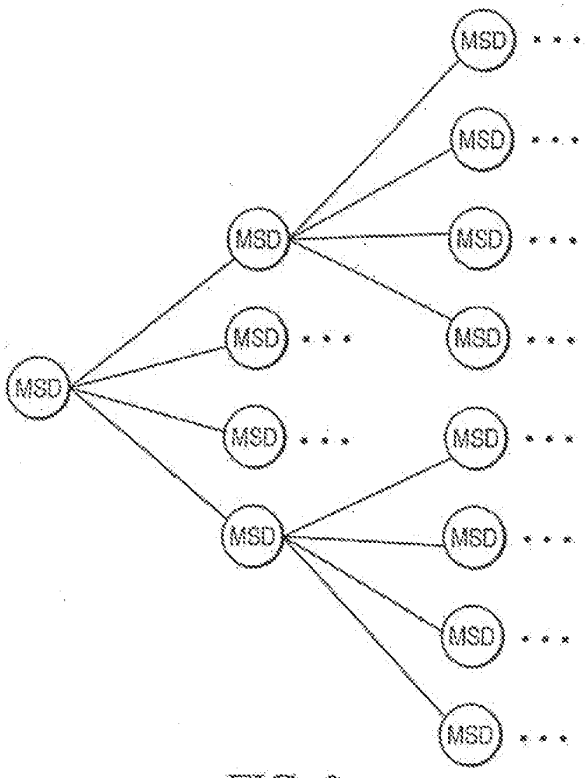

MSD

FIG. 9

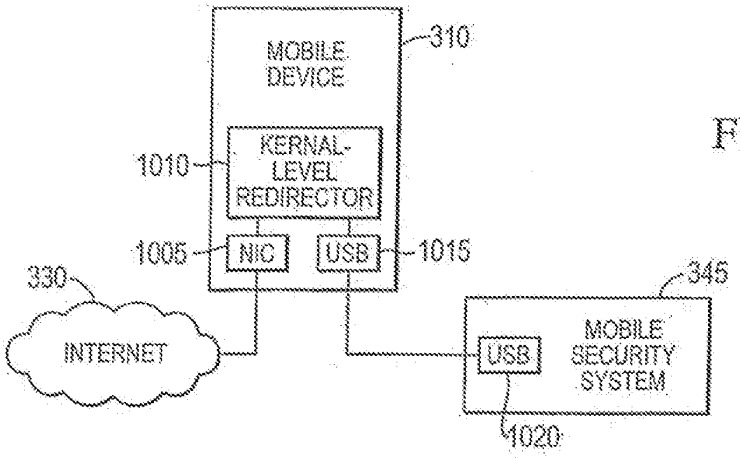


FIG. 10A

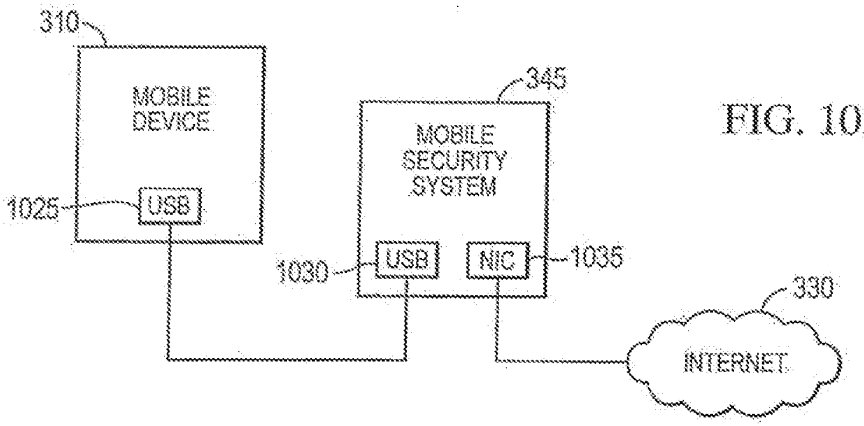


FIG. 10B

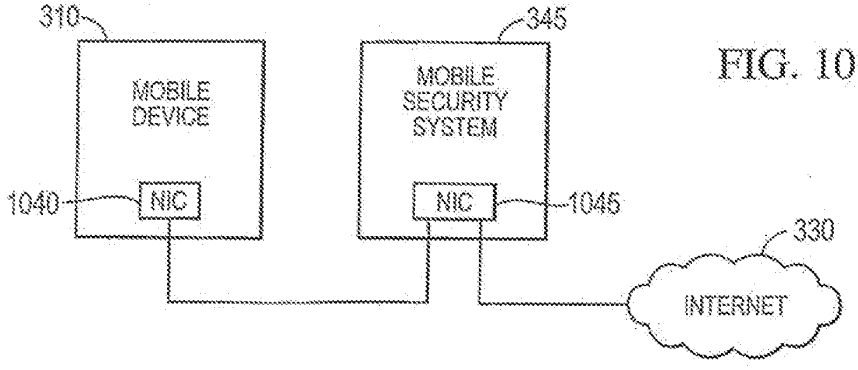


FIG. 10C

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 June 2008 (05.06.2008)

PCT

(10) International Publication Number
WO 2008/066419 A1

(51) International Patent Classification:
H04Q 3/00 (2006.01) *H04L 12/24* (2006.01)

(21) International Application Number:
PCT/SE2006/001364

(22) International Filing Date:
29 November 2006 (29.11.2006)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)** [SE/SE]; S-164 83 Stockholm (SE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LIDSTRÖM, Mattias** [SE/SE]; Kungsholms Kyrkoplan 3A, S-112 24 Stockholm (SE). **MATTI, Mona** [SE/SE]; Danielsvägen 2A, S-131 40 Nacka (SE). **KVERNVIK, Tor** [NO/SE]; Ropgränd 8, S-187 42 Täby (SE). **BJÖRK, Niklas** [SE/SE]; Krokivägen 13, S-175 47 Järfälla (SE). **LARSSON, Tony** [SE/SE]; Nils Åhlins väg 2, S-194 76 Upplands Väsby (SE). **BENITO GARCIA, Raul** [ES/ES]; C/Cordoba 7-4a, ES-28770 Colmenar Viejo (ES).

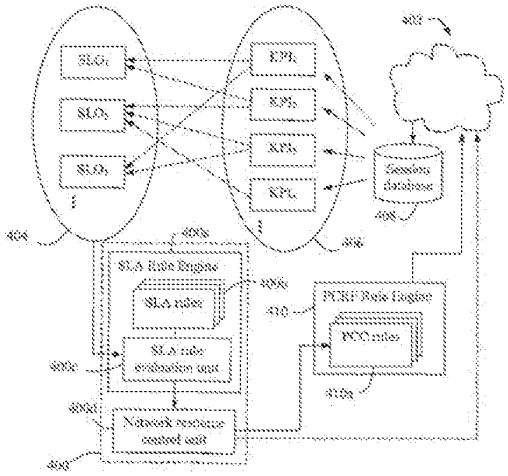
(74) Agent: **RAGSTÖM, Hans**; Bergenstråhle & Lindvall AB, Box 17704, S-118 93 Stockholm (SE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) Title: A METHOD AND ARRANGEMENT FOR CONTROLLING SERVICE LEVEL AGREEMENTS IN A MOBILE NETWORK.



(57) Abstract: A method and an arrangement in a service assurance module (400) for controlling service level agreements for communication services in a mobile services access network (402). A network performance rule (400b) is created for a service level agreement, and is based on an associated performance indicator (406) indicating the current network performance with respect to the fulfilment of the service level agreement. The performance indicator is then monitored, and the network performance rule is evaluated on a regular basis based on the performance indicator. If the network performance rule is not satisfied, an agreement violation is registered and/or measures are taken to improve the network performance.

WO 2008/066419 A1

A METHOD AND ARRANGEMENT FOR CONTROLLING SERVICE LEVEL AGREEMENTS IN A MOBILE NETWORK.

TECHNICAL FIELD

The present invention relates generally to a method and arrangement for determining whether service level agreements are fulfilled in a mobile services access network. In particular, any quality-related problems in the network that may have a negative impact on service level agreements, may be detected and redeemed.

BACKGROUND

With the emergence of 3G mobile telephony, new packet-based communication technologies using IP (Internet Protocol) have been developed to support wireless communication of multimedia. For example, communication protocols in GPRS (General Packet Radio Service) and WCDMA (Wideband Code Division Multiple Access) support packet-switched multimedia services, as well as traditional circuit-switched voice calls.

A network architecture called "IP Multimedia Subsystem" (IMS) has been developed by the 3rd Generation Partnership Project (3GPP) as a platform for handling multimedia services and sessions in the packet domain, based on IP transport. Thus, an IMS network can be used to initiate and control multimedia sessions for IP enabled terminals being connected to any type of access networks. A signalling protocol called "SIP" (Session Initiation Protocol, according to the standard IETF RFC 3261) has been defined for handling sessions in IMS networks. A "SIP-enabled" terminal can thus use this standard to initiate and

terminate multimedia communications by means of its home IMS network.

Fig. 1 is a simplified schematic illustration of a basic network structure for providing multimedia services for a mobile terminal A by means of an IMS network 100. The terminal A is connected to a mobile access network 102 and communicates in a multimedia session with another terminal B which may be connected to the same access network 102 or another access network (not shown). Alternatively, terminal A may communicate with a content server or the like, e.g. for downloading some multimedia content therefrom.

The access network 102 is connected to IMS network 100 which is the "home" IMS network of terminal A and therefore handles the session with respect to terminal A. The same IMS network 100 or another one (not shown), may handle the session on behalf of terminal B. Moreover, multimedia services are always handled by the terminal's home IMS network even if terminal A would be roaming in a visited access network.

The session is controlled by specific nodes in the IMS network 100, here generally referred to as "session managing nodes" 104. These nodes include S-CSCF (Serving Call Session Control Function), I-CSCF (Interrogating Call Session Control Function) and P-CSCF (Proxy Call Session Control Function, according to the conventional IMS architecture. Briefly described, a P-CSCF node acts as an entry point towards the IMS network 102 from access networks, including network 100, a plurality of S-CSCF nodes are assigned to active terminals for managing their sessions using SIP signalling, and an I-CSCF node acts as a gateway for SIP messages from other IMS networks.

The IMS network 102 also includes one or more application servers 106 for enabling various multimedia services, and a main database element HSS (Home Subscriber Server) 108 containing subscriber and authentication data. The various functions of the shown network elements 104-108 are generally known in the art, and not necessary to describe here further to understand the context of the present invention. Of course, IMS network 102 contains various other nodes and functions not shown here for the sake of simplicity. In the figure, the thick two-way arrow illustrates the communication of payload data between the two terminals A, B. The thin two-way arrow illustrates the communication of control messages, typically according to SIP in the case of IMS.

A policy unit 110 associated with the access network 102 is also connected to the IMS network 100. The policy unit 110 is basically responsible for authorising communication sessions for terminals connected to the access network 102, based on various predetermined policies and rules. The policy unit 110 generally operates according to a function implemented therein called PCRF (Policy and Charging Rule Function), sometimes alternatively referred to as PDF (Policy Decision Function).

The PCRF uses a rule engine that evaluates session information related to a requested session, in order to determine whether the session can be allowed or not. A session may be allowed conditionally, e.g., "session X is allowed if condition Y is met". Session admission rules defined in the rule engine of policy unit 110 specify when to allow sessions for specific users and/or services. The rules can be combined into policies which may be defined for specific users or services, or both. When a session request

is received from a user for a specific service, a relevant policy or rule combination is thus selected and evaluated for the session request for allowance or rejection.

Many mobile applications and services require a certain Quality of Service QoS in order to provide a satisfying and expected result to end-users. For UMTS networks, four main traffic classes have been defined: "conversational class", "streaming class", "interactive class" and "background class", in order to classify different needs and requirements regarding bit rates and delays. These traffic classes are primarily distinguished by their requirements regarding transfer delays, such that applications of the conversational class tolerate only small delays, sometimes also referred to as "real-time", whereas the background class is applied to the least delay-sensitive applications, sometimes also referred to as "best effort".

The selection of a UMTS traffic class for an application is used for assigning a suitable physical channel in the mobile network, generally referred to as a RAB (Radio Access Bearer), in order to optimise the scarce radio recourses and other transport resources in the radio access and core networks whilst maintaining an acceptable and expected level of quality for the end-user.

Another factor that may dictate the QoS requirements is the subscription type. Thus, subscribers may be offered specific service levels with respect to QoS for different subscription types. For example, using current terms, a "Platinum" subscription may be defined to guarantee a relatively high level of service or QoS, whereas "Gold", "Silver" and "Bronze" subscriptions may provide successively lower levels of service or QoS. Moreover, specific services

may also be offered with different levels of QoS, e.g. at different prices.

Fig. 2 illustrates a conventional process when a mobile terminal A will communicate multimedia with another party B, involving different **stages 2:1 - 2:9** as illustrated. Here, each stage may generally represent the communication of various messages and/or data back and forth which are not necessary to describe in more detail.

IMS nodes P-CSCF 202 and S-CSCF 204 are shown here in a schematic "control plane" for control signalling over the IMS network. Further, a gateway node GW 200 in the core part of a mobile network, typically a GGSN (Gateway GPRS Switching Node), is shown in a schematic "bearer plane" used for the data transport. The gateway node GW 200 is connected to an associated policy unit 206, here denoted PCRF, which is thus basically responsible for authorising communication sessions to terminals in the associated mobile network, as described above. In some systems, a so-called "Policy and Charging Enforcement Function" (PCEF) in the GGSN node typically communicates with the PCRF node over a Gx interface to control the access admission, and the PCRF node communicates with the P-CSCF node over an Rx interface.

In a first **stage 2:1** of the illustrated procedure, terminal A obtains an initial connection with the mobile network, typically involving the activation of a PDP context and a Radio Access Bearer RAB, as established by the gateway node GW 200. A particular PCRF 206 is also selected for terminal A by GW 200 at this point. The established access bearer is primarily used for carrying occasional control messages of minor size such as service requests, allowing for relatively low bandwidth and fairly long delays.

In a next **stage 2:2**, GW 200 sends relevant access information to PCRF 206, including a subscription identifier and the bearer information established in stage 2:1. Thereby, a "state" is created in PCRF 206 for a "bearer session", meaning that the received access information for terminal A is retained in PCRF 206. In this stage, PCRF 206 also responds to GW 200 by sending a so-called "basic charging rule" that has been selected or created based on the received access information. The basic charging rule includes a charging key, a charging identity and any service data flow filters that should be applied on communicated data according to that rule. A certain basic level of QoS, typically the "best effort" level, has now also been established for non-IMS data services such as the communication of various control messages as mentioned above.

When a user activates a selected application in terminal A in order to communicate with terminal B in this case, an SIP-based user agent ("SIP UA") in terminal A performs a session negotiation in a following **stage 2:3**, which may involve the exchange of terminal capabilities with terminal B. In the session negotiation, terminal A may start by sending an SIP INVITE message as a session request towards terminal B containing session-specific parameters typically including a proposed coder/decoder (or "codec") for the session.

After the session negotiation, as the session parameters have been settled, the P-CSCF node 202 sends corresponding service information to PCRF 206, in a next **stage 2:4**. Thereby, another state is created in PCRF 206 for a "service session", meaning that the received service information for the session of terminal A is retained in

PCRF 206. Moreover, PCRF 206 determines whether the requested session should be allowed or not by applying a suitable policy.

If PCRF 206 can allow the requested session according to the applied policy for the given user and service, an OK message is sent to the P-CSCF node 202 which also sends an OK message to the terminal A, in a **stage 2:5**. The application activated in terminal A then requests a bearer from a bearer part in terminal A, in a following **stage 2:6**, in order to satisfy QoS requirements for communication according to the application. Terminal A then issues a bearer request for service data transport to GW 200 in the mobile network, in a next **stage 2:7**. Alternatively, the establishment of bearer may also be initiated by the mobile network.

In a further **stage 2:8**, GW 200 now establishes a new bearer for the forthcoming data transport, as required by the service, and sends relevant QoS information on the established service bearer to PCRF 206 where the bearer session state is updated accordingly. In this stage, PCRF 206 also responds to GW 200 by sending an updated charging rule based on the received service and QoS information. The updated charging rule includes a service identity, a charging identity and any service data flow filters according to that rule. Hence, a certain level of QoS has also been established for the requested service, typically a higher QoS level than the one established for non-IMS data services in stage 2:2 above.

The above-mentioned basic charging rule and updated charging rule established for terminal A in stages 2:2 and 2:8, respectively, are often referred to as "PCC (Policy and Charging Control) rules", in effect being rules for session

admission. In general, a PCC rule is used for detecting packets belonging to a service data flow, identifying the service, and for providing relevant charging parameters and policy control for a data flow. A PCC rule may be predefined or created dynamically, and includes basically a rule name, a service identifier, service data flow filters, QoS parameters and charging parameters.

Finally, terminal A can start the session involving the communication of data between the two terminals A and B, as illustrated by a final **step 2:9**. The new service bearer or RAB established in stage 2:8 should be more stable and reliable than the access bearer established for control messages in stage 2:2, to provide a "guaranteed" level of QoS according to the service involved as controlled by the updated charging rule. As mentioned above, different levels of QoS may also be expected depending on the subscription and/or selected level of service.

However, the mobile network will undoubtedly sometimes fail to fulfil the agreed level of QoS during such sessions. The reasons for not providing the agreed levels of service can be many, often depending on the current traffic situation such as cell congestion and interference problems, but also on factors related to the operation of the network such as inadequate power regulation, equipment failure, inaccurate channel allocation and poor radio coverage, etc.

Today, there is no efficient way to monitor service level agreements in mobile networks, and consequently it cannot be controlled that these agreements are met properly. Hence, if a customer complains that an expected QoS level has not been fulfilled during a session, the network operator has no means to check the relevance thereof. Moreover, the operator cannot work proactively to ensure

that service level agreements are fulfilled. Instead, a reactive process must be used for investigating a network in order to identify such problems and where they typically occur. This is an expensive and time consuming activity that nonetheless must be undertaken in order to improve the fulfilment of service agreements, and to avoid customer dissatisfaction and the paying of resulting penalty fees.

SUMMARY

It is an object of the present invention to address the problems outlined above. More specifically, it is an object of the present invention to make it possible to detect and identify any performance shortcomings that may jeopardise service level agreements in a mobile services access network, in order to take suitable measures to remedy such deficiencies.

These objects and others can be obtained by providing a method and an arrangement according to the independent claims attached below.

According to one aspect, the present invention involves a method of controlling a service level agreement for a communication service in a mobile services access network. The method includes creating a network performance rule valid for the service level agreement and being based on an associated performance indicator indicating the current network performance with respect to the fulfilment of the service level agreement. The method further includes monitoring the performance indicator and evaluating the network performance rule based on the monitored performance indicator. An agreement violation is then registered and/or measures are taken to improve the network performance, if the evaluated network performance rule is not satisfied.

According to another aspect, the present invention also involves a service assurance module in a mobile services access network for controlling service level agreements for communication services. The service assurance module comprises means for creating a network performance rule valid for a service level agreement and being based on an associated performance indicator indicating the current network performance with respect to the fulfilment of the service level agreement. The service assurance module further comprises means for monitoring the performance indicator, means for evaluating the network performance rule based on the monitored performance indicator, and means for registering an agreement violation and/or for taking measures to improve the network performance, if the network performance rule is not satisfied.

The invented method and service assurance module may optionally have the following further characteristics.

The measures to improve network performance may include modifying or adding a session admission rule applied for individual terminals and session requests, to introduce access constraints. The session admission rule is then used by a rule engine of a policy unit to control the admission of individual sessions.

The network performance rule may include a degradation rule specifying a degradation level of the monitored performance indicator, and a violation rule specifying a violation level of the monitored performance indicator when the service level agreement is deemed to be violated. The degradation and violation rules are conditional by dictating said measures to improve network performance if the respective rule is not met, and the

violation rule is evaluated if the degradation rule is not satisfied.

The network performance rule may be dependent on a plurality of associated performance indicators indicating different aspects of the current network performance with respect to the fulfilment of the service level agreement.

The monitored performance indicator(s) may be derived from a session database used by the mobile network and may relate to any of: bandwidth occupation, latency, packet loss rate, available bitrate, or network and service accessibility.

The present invention makes it possible to create specific service level agreements for individual services and/or subscribers, and to ensure that these agreements can be substantially satisfied. Performance problems occurring in a network can also be identified and located, by means of monitored performance indicators declared in network performance rules. Thereby, the network operator can work proactively in order to avoid the violation of service level agreements.

Further features of the present invention and its benefits will be explained in the detailed description below.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described in more detail by means of preferred embodiments and with reference to the accompanying drawings, in which:

- Fig. 1 is a schematic communication scenario of a multimedia session between two mobile terminals A and B using a conventional IMS network, according to the prior art.

- Fig. 2 illustrates a conventional signalling flow during the establishment of a session between terminals A and B using a policy unit PCRF, according to the prior art.
- Fig. 3 is a diagram illustrating the fluctuations of a monitored network performance indicator over time.
- Fig. 4 is a block diagram illustrating the operation of a network performance rule engine connected to a session admission rule engine, according to one embodiment.
- Fig. 5 is a flow chart illustrating a procedure for overseeing service level agreements in a mobile services access network, according to another embodiment.
- Fig. 6 is a flow chart illustrating a procedure for controlling network resources in a mobile services access network based on at least one monitored performance indicator, according to another embodiment.

DETAILED DESCRIPTION

Briefly described, the present invention involves a service assurance module or similar including a "network performance rule engine", referred to as the "SLA rule engine" for short in the following description, which can be used to improve the fulfilment of service level agreements (SLA) in a mobile services access network. The SLA rule engine is thus used for monitoring whether such agreements are satisfied, based on network performance rules that have been derived or translated from the agreements, and can also be used for controlling network resources to remedy or avoid any violations thereof.

The SLA rule engine is connected to the above-described session admission rule engine in the policy unit PCRF, hereafter referred to as the "PCRF rule engine". Thus, whereas the session admission rules in PCRF (i.e. PCC rules)

are applied for individual terminals and session requests to control session admission, the network performance rules in the SLA rule engine are evaluated with respect to network performance for subscriber groups and/or specific services, which in turn depends on the outcome from the PCRF rule engine. Hence, the results from the SLA rule engine indicate the extent of service level agreement fulfilment. The outcome from the SLA rule engine can then be used for detecting any overloading or deficiencies in the mobile network, and for improving the network performance by controlling network resources, in a manner that will be described below by means of exemplary embodiments.

The rules in the SLA rule engine may be frequently evaluated on a regular basis, e.g. at intervals of five minutes. An SLA rule may specify a maximum period of time during which an SLA violation can be tolerated, which of course will depend on the nature of agreement and service. In other words, in may be deemed acceptable that an SLA is violated temporarily, as long as it is fulfilled again within the set time period.

In order to determine whether an SLA is fulfilled or not, one or more performance indicators are monitored, hereafter referred to as the "Key Performance Indicator", KPI. A KPI generally indicates the performance in the network with respect to a predefined factor or parameter relevant for an evaluated SLA rule, that can be detected or measured to determine a current KPI value or level, which typically fluctuates over time.

The KPI values may be derived from information in a session database maintained in the mobile network, but also from other information sources, and the present invention is not limited in this respect. For example, a KPI

may relate to bandwidth occupation, latency, packet loss rate, available bitrate, or network and service accessibility, etc., although the present invention is not limited to these examples neither. Thus, a plurality of KPI's related to different aspects of network performance may be monitored in order to evaluate rules in the SLA engine and ultimately the fulfilment of service level agreements.

Fig. 3 is a schematic diagram of an exemplary performance indicator KPI being monitored, where the horizontal axis represents time and the vertical axis represents the KPI level. In this example, the KPI is related to the occupation of bandwidth in a certain area such as a cell or group of cells in a mobile network, and/or for a specific service. The curve 300 thus represents the measured KPI over time. In general, the KPI level is high when a large percentage of the available bandwidth is occupied, and vice versa.

In this case, a violation level of the KPI has been predefined, as indicated in the figure. Basically, a service level agreement SLA ensuring a certain amount of free bandwidth in the monitored area and/or service, will be violated if the KPI value exceeds the violation limit for more than a predetermined maximum acceptable time period T. Further, a relatively lower KPI level has also been defined over which the bandwidth availability is considered to be degraded, as indicated in the figure.

As can be seen in the figure, the fluctuating KPI stays below the degradation level until the time t_1. Thereafter, the KPI exceeds the degradation level, and the KPI increases further to even exceed the violation level at the time t_2. In this case, the violation level has been

exceeded for more than the maximum acceptable time period T after the time t_3, until the KPI decreases to fall below the violation level at the time t_4, and finally also below the degradation level at the time t_5. Hence, a service level agreement being dependent on the KPI depicted in Fig. 3 is considered to be violated during the period between t_3 and t_4, i.e. after the maximum acceptable time period T.

In the example above, as soon as the KPI level rises above the degradation level after t_2, it is possible to determine why and where the degradation has occurred, as indicated by the KPI. Thereby, suitable measures can be taken in the mobile network at some point after t_2 to control network resources, in order to remedy the problem and improve the bandwidth availability such that the KPI level does not rise further to unacceptable levels. For example, the KPI results can be used for modifying and/or adding selected PCC rules to control the bandwidth occupation, e.g., by introducing access constraints or the like in the PCC rules. Thus, once the KPI becomes degraded, session admission may be denied or restricted for certain services and/or subscribers in order to improve the bandwidth availability for other services and/or subscribers with higher priority.

In this way, the performance problem can be redeemed proactively by controlling network resources, such that the KPI level is kept within acceptable limits, as indicated by an exemplary dashed curve 302 in Fig. 3. Violation of the associated SLA can then be avoided or at least reduced.

In order to evaluate the fulfilment of service level agreements in a mobile network, different "Service Level Objectives" (SLO) may be defined based on monitored

performance indicators (KPI), such that one or more specific performance indicators are associated with each service level objective. Thus, a service level objective SLO is basically deemed to be fulfilled as long as no associated performance indicator KPI exceeds a predefined violation level for more than a maximum acceptable time period.

Fig. 4 illustrates schematically the operation of a service assurance module 400 comprising an SLA rule engine 400a, according to one embodiment. In the SLA rule engine 400a, one or more SLA rules 400b (i.e. network performance rules) have been defined to evaluate the network performance with respect to SLA's (i.e., service level agreements) for subscribers and/or services in a mobile services access network 402.

A plurality of service level objectives 404 have also been defined for the SLA rule engine 400a, each being dependent on one or more performance indicators 406. An SLA rule may basically involve one or more SLO's. In this example, a first service level objective SLO_1 is associated with performance indicators KPI_1 and KPI_2. Further, a second service level objective SLO_2 is associated with performance indicators KPI_2, KPI_3 and KPI_4, and a third service level objective SLO_3 is associated with performance indicators KPI_3 and KPI_4, respectively.

The SLA rule engine 400 further includes a rule evaluation unit 400c receiving KPI measurements for each SLO 404, which is adapted to evaluate the SLA rules 400b based on the received KPI measurements. The mobile network 402 and the subscribers and services therein effectively constitute the environment in which the measured KPI's are formed. The KPI's 406 may be derived from a session database 408 used by

the network 402, as illustrated by arrows, but also from other sources of network information, not shown.

The service assurance module 400 further comprises a resource control unit 400d adapted to generally control network resources, in response to the evaluation of the SLA rules in the evaluation unit 400c. A conventional PCRF rule engine 410 is used by the mobile network 402 to control session admission for its subscribers and services, and the network resource control unit 400d is connected thereto. The control unit 400d is further adapted to modify and/or add PCC rules 410a in the PCRF rule engine 410 in response to the SLA rule evaluation, to control resources in network 402 such that any SLA violations can be avoided or at least reduced, as described above.

The control unit 400d may also be adapted to influence resources and operations in network 402 in other ways, as schematically indicated by the dashed arrow. For example, channel allocation schemes, power regulation and radio coverage, etc., may be modified in response to the evaluation of SLA rules using methods and routines well-known to the skilled person.

It should be noted that the arrangement depicted in Fig. 4 is purely logical and can be implemented in different ways. For example, the SLA rule engine 400a and the network resource control unit 400d can be arranged together as the shown service assurance module 400, but may also be arranged as separate connected entities. The components 400a-d in the shown service assurance module 400 may further be integrated with the PCRF Rule Engine 410, or implemented as a separate connected module, node or system. An Operation and Maintenance (O/M) interface (not shown) to the SLA rule engine 400a may be used allowing an Operation

and Support System (OSS) to supervise the fulfilment of SLA's, and providing the possibility to operators in a Network Operation Centre (NOC), to modify parameters in SLA Policies to control SLA fulfilment.

As shown by means of various arrows, the arrangement in Fig. 4 basically produces a network control "loop" for optimising the usage of network resources in order to fulfil prevailing SLA's. Thus, the session database 408 (and possibly also other information sources, not shown) provides KPI values 406, associated with different SLO's 404. The evaluation unit 400c evaluates the SLA rules based on the obtained KPI values and the associated SLO's, and the outcome is used by the network resource control unit 400d to modify and/or add PCRF rules 410a (and possibly also to influence the network 402 in other ways, as indicated by the dashed arrow), if necessary. As a result, the network 402 will provide a different performance, as reflected in the session database 408 and otherwise, such that the measured KPI's 406 hopefully become improved.

Fig. 5 is a flow chart of steps executed by a service assurance module or similar, adapted to oversee service level agreements in a mobile services access network and control network resources such that violations of the service level agreements can be avoided or at least reduced. In a first **step 500**, a network performance rule is created based on one or more service level agreements valid for the mobile network. At least one associated performance indicator is monitored in a next **step 502**, the level of which indicating whether any of the created service level agreements is fulfilled or not.

In a following **step 504**, it is determined whether the network performance rule is satisfied or not, based on

the monitored performance indicator(s). If the rule is satisfied, the process returns to step 502, such that steps 502 and 504 are repeated on a regular basis, e.g. at intervals of five minutes.

However, if the network performance rule is not satisfied in step 504, some different options are possible for the next **step 506**. One option is that a failure or breach of the service level agreement associated with the created network performance rule can be simply registered or logged in a database or the like, e.g. in order to collect SLA statistics. Another option in step 506 is that suitable measures are taken to somehow improve the network performance as relevant to the evaluated network performance rule, in response to a negative outcome of step 504, in order to influence the environment of the mobile network in a way that the service level agreement can be safely met again. For example, at least one session admission rule can be modified or added in a PCRF rule engine or similar, e.g. as described for Fig. 4 above. As indicated in step 506, either or both alternatives are possible.

As mentioned above in connection with Fig. 3, two different network performance rules may be set for a particular performance indicator, which can be used for controlling network resources in an mobile network to avoid or reduce any violation of the corresponding SLA. A first "degradation" rule specifies a degradation level or limit of the KPI as described for Fig. 3. The degradation rule is also "conditional" by dictating measures to be taken for changing the admission of sessions in order to remedy the network performance degradation, if the rule is not met. Further, a second "violation" rule specifies a violation level or limit of the KPI when the network performance has

deteriorated so that a service level agreement is deemed to be violated, e.g. when the violation limit is exceeded for more than a predetermined time period. The violation rule is also "conditional" in the same manner.

Fig. 6 is a flow chart illustrating a procedure for controlling network resources proactively in a mobile services access network, according to another embodiment, by using the above-mentioned degradation and violation rules created for a prevailing SLA, to avoid or at least reduce the violation of that SLA. The following example is concerned with only one performance indicator KPI and one session admission rule, even though the present invention is not limited in this respect but can be used for any number of performance indicators, session admission rules, network performance rules and service level agreements.

It is assumed that a degradation rule and a violation rule has been created with respect to a certain SLA as described above. In a first illustrated **step 600**, a performance indicator KPI is monitored, being relevant for the two network performance rules. In a next **step 602**, the degradation rule is evaluated to determine whether it is met or not, according to the monitored performance indicator. As long as the degradation rule is satisfied, it is still evaluated on a regular basis as described above, thus repeating steps 600 and 602 frequently.

However, if the degradation rule is not satisfied in step 602, a session admission rule is modified or added, in a **step 604**, hopefully impacting the monitored performance indicator. The session admission rule (i.e. PCC rule) may reside in a PCRF rule engine implemented in a policy unit controlling session admission for individual session requests in the mobile network, as shown in Fig. 4. The

degradation rule may thus dictate that if the KPI exceeds the degradation level or limit, the session admission rule should be modified in a specific way. For example, if the monitored performance indicator is bandwidth usage for service X, the degradation rule may dictate: "if the allocated bandwidth for service X exceeds 60%, downgrade bandwidth 1 level for Bronze subscribers".

Once the session admission rule has been modified or added, the performance indicator hopefully changes to the better, which is monitored further in a following **step 606**. It is then determined whether the violation rule is satisfied in a further **step 608**. If so, the process may return to the initial steps 600 and 602 to find out whether the performance indicator has improved such that the degradation rule is now also met.

However, if neither the violation rule is satisfied in step 608, an SLA violation may be registered or logged in a suitable database or the like, as indicated by an optional **step 610**. In any case, a session admission rule is once again modified or added, by returning to step 604. At this stage, the session admission rule modified or added previously in step 604 may be further modified this time, e.g. by introducing harder access constraints or the like. However, any other session admission rule could be modified or added, depending on the implementation, and the present invention is not limited in this respect.

The violation rule evaluated in step 608 may thus dictate that if the KPI exceeds the violation level, the previously modified or added session admission rule should be modified further. For example, if the degradation rule example given above has been used for steps 602 and 604, the violation rule may dictate: "if the allocated bandwidth for

service X exceeds 80% for more than 5 minutes, deny access for Bronze subscribers".

It is possible to define network performance rules in many different ways, based on various performance indicators. As mentioned above, conditional degradation and violation rules may dictate different measures to be taken when the rules are not met, in order to remedy or avoid the network performance problems. As mentioned above, these measures may include measures taken on various network operational tools and methods, but involve primarily the modification or addition of session admission rules which will typically produce a swift result. For example, a frequently used service with short lived sessions can be targeted for enforcing restrictions since the effects will be instant in that case.

When it is determined that a conditional network performance rule is met again, after having introduced a restriction by modifying or adding a session admission rule according to the above, the situation may be restored by removing the restriction, e.g. after a predetermined delay period and/or when the monitored KPI falls below a certain level. The evaluated network performance rule may specify how and when a previously introduced restriction is to be removed.

More or less complex rule schemes may be created where different services affect each other. For example, a conditional degradation rule may dictate that if the allocated bandwidth for service X exceeds a certain limit, agreements on services X, Y and Z cannot be fulfilled for Platinum users, and therefore the access to one or more of these services is restricted for Bronze users.

Different service level agreements between a network operator and subscribers will be given different priorities, and different performance indicators KPI's can be given different priorities within each service level agreement as well. Using the present invention, it is therefore possible to create specific service level agreements for individual services and subscribers or subscriber groups, also ensuring that the agreements can be satisfied. The invention also provides a solution to identify any performance problems in a network and where they occur, since each KPI declared in the network performance rules can be monitored easily and efficiently. The network operator can thus work proactively in order to enhance service offerings and ultimately the end user experience.

The present invention may be implemented and configured as follows: SLA rules for one or more subscriber groups or categories may be created in connection with introducing a service. The service will thus have policies that are derived from the SLA which may be offered along with the service. These policies are made up of session admission rules (PCC rules) set in the PCRF rule engine and network performance rules (SLA rules) set in the SLA rule engine.

The PCC rules are focused on the single user and dictates how well a service will perform for the given user. On the other hand, the SLA rules are focused on how the service performs in the network, e.g. for a certain subscriber group or class, and how well the service performance matches the SLA between a service provider and an operator. The SLA rules are generally fewer than the PCC rules and require little effort from the operator of the

system to install. Since an SLA rule is valid for a subscriber group (i.e., a large aggregated group of end-users), the number of such rules will be limited.

The fact that the PCC and SLA rules focus somewhat differently allows for different outcomes when changing either of the two. When the SLA rules operate on the PCRF policies by impacting the PCC rules as described above, both QoS and SLA fulfilment can be assured. Initially, a network operator would probably prefer to use the SLA rule engine only for "important" customers that actually have SLA's with the operator such as corporate customers and services. The present invention also enables the possibility for an operator to use the SLA rule engine to provide important feedback for subscribers and services based on various groupings that the operator is free to define.

Thus, the following useful characteristics may generally be obtained when using the present invention:

- An SLA can be expressed in terms of concrete rules.
- An SLA and the service performance can be easily but carefully monitored by frequently evaluating the corresponding SLA rule.
- The network performance and the end user experience can be enhanced proactively.
- The operator's revenue can hopefully be increased.
- The amount of incoming complaints to a service/support desk (helpdesk) or the like can be reduced.

Moreover, As soon as the SLA rule engine outcome gives rise to a change of existing admission rules for any

subscriber category or service, the helpdesk can be notified in runtime about the decision taken. For example, the SLA rule engine may provide the message that "all Bronze subscribers for the free Poker Service are currently at reject status due to extremely high interest in the Platinum World Cup Soccer Service". This type of notifications will thus enable improved end-user support through the helpdesk.

While the invention has been described with reference to specific exemplary embodiments, the description is in general only intended to illustrate the inventive concept and should not be taken as limiting the scope of the invention. For example, the SIP signalling protocol and IMS concept have been used throughout when describing the above embodiments, although any other standards and service networks may basically be used. The present invention is defined by the appended claims.

CLAIMS

1. A method of controlling a service level agreement for a communication service in a mobile services access network (402), comprising the following steps:
- creating a network performance rule (400b) valid for said service level agreement and being based on an associated performance indicator (406) indicating the current network performance with respect to the fulfilment of said service level agreement,
- monitoring said performance indicator,
- evaluating said network performance rule based on the monitored performance indicator, and
- registering an agreement violation and/or taking measures to improve the network performance, if the evaluated network performance rule is not satisfied.

2. A method according to claim 1, wherein said measures to improve network performance includes modifying or adding a session admission rule applied for individual terminals and session requests, to introduce access constraints.

3. A method according to claim 2, wherein the session admission rule is used by a rule engine (410) of a policy unit to control the admission of individual sessions.

4. A method according to any of claims 1-3, wherein the network performance rule includes a degradation rule specifying a degradation level or limit of the monitored performance indicator, and a violation rule specifying a violation level or limit of the monitored performance indicator when the network performance has deteriorated

so that said service level agreement is deemed to be violated.

5. A method according to claim 4, wherein the degradation and violation rules are conditional by dictating said measures to improve network performance if the respective rule is not met.

6. A method according to claim 5, wherein the degradation rule is evaluated on a regular basis and the violation rule is evaluated if the degradation rule is not satisfied.

7. A method according to any of claims 4-6, wherein the service level agreement is deemed to be violated if the monitored performance indicator exceeds the violation level for more than a maximum acceptable time period.

8. A method according to any of claims 1-7, wherein the network performance rule is dependent on a plurality of associated performance indicators indicating different aspects of the current network performance with respect to the fulfilment of said service level agreement.

9. A method according to any of claims 1-8, wherein a Service Level Objective is defined based on one or more monitored performance indicators, such that the service level objective is deemed to be fulfilled as long as the associated performance indicators remain within acceptable limits.

10. A method according to any of claims 1-9, wherein said one or more monitored performance indicators relate to any of: bandwidth occupation, latency, packet loss rate, available bitrate, or network and service accessibility.

11. A method according to any of claims 1-10, wherein the monitored performance indicator is derived from a session database (408) used by the mobile network.

12. A service assurance module (400) in a mobile services access network (402) for controlling service level agreements for communication services, comprising:
- means for creating a network performance rule (400b) valid for a service level agreement and being based on an associated performance indicator (406) indicating the current network performance with respect to the fulfilment of said service level agreement,
- means for monitoring said performance indicator,
- means for evaluating said network performance rule based on the monitored performance indicator, and
- means for registering an agreement violation and/or for taking measures to improve the network performance, if the evaluated network performance rule is not satisfied.

13. A module according to claim 12, wherein said measures taking means is adapted to modify or add a session admission rule applied for individual terminals and session requests, to introduce access constraints.

14. A module according to claim 13, wherein the session admission rule is used by a rule engine (410) of a policy unit to control the admission of individual sessions.

15. A module according to any of claims 12-14, wherein the network performance rule includes a degradation rule specifying a degradation level or limit of the monitored performance indicator, and a violation rule specifying a violation level or limit of the monitored performance indicator when the network performance has deteriorated so that said service level agreement is deemed to be violated.

16. A module according to claim 15, wherein the degradation and violation rules are conditional by dictating said measures to improve network performance if the respective rule is not met.

17. A module according to claim 16, wherein the evaluating means is adapted to evaluate the degradation rule on a regular basis, and to evaluate the violation rule if the degradation rule is not satisfied.

18. A module according to any of claims 15-17, wherein the service level agreement is deemed to be violated if the monitored performance indicator exceeds the violation level for more than a maximum acceptable time period.

19. A module according to any of claims 12-18, wherein the network performance rule is dependent on a plurality of associated performance indicators indicating different aspects of the current network performance with respect to the fulfilment of said service level agreement.

20. A module according to any of claims 12-19, further comprising means for defining a Service Level Objective based on one or more monitored performance indicators, such that the service level objective is deemed to be fulfilled as long as the associated performance indicators remain within acceptable limits.

21. A module according to any of claims 12-20, wherein said one or more monitored performance indicators relate to any of: bandwidth occupation, latency, packet loss rate, available bitrate, or network and service accessibility.

22. A module according to any of claims 12-21, wherein the monitoring means is adapted to derive the performance indicator from a session database (408) used by the mobile network.

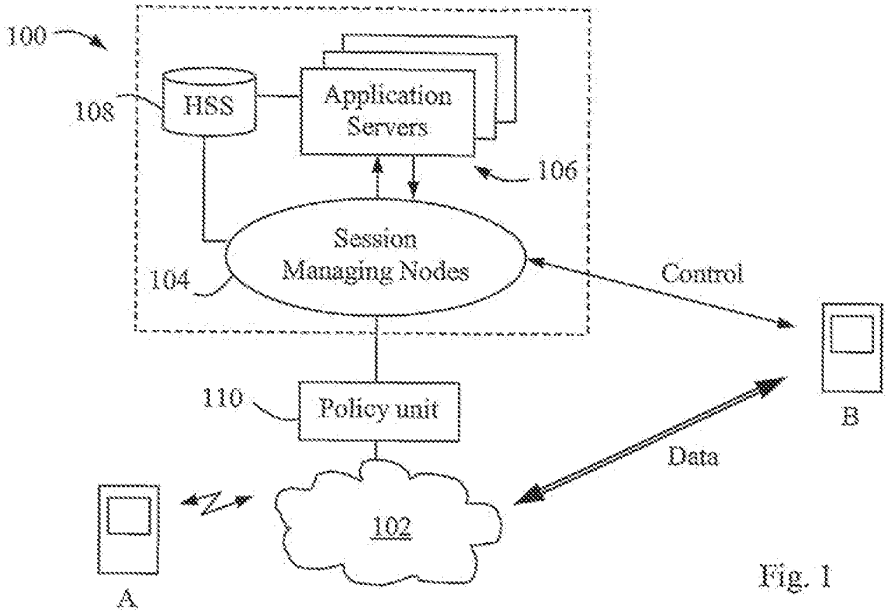

100
108
HSS
Application Servers
106
Session Managing Nodes
104
Control
110
Policy unit
B
Data
102
A

Fig. 1

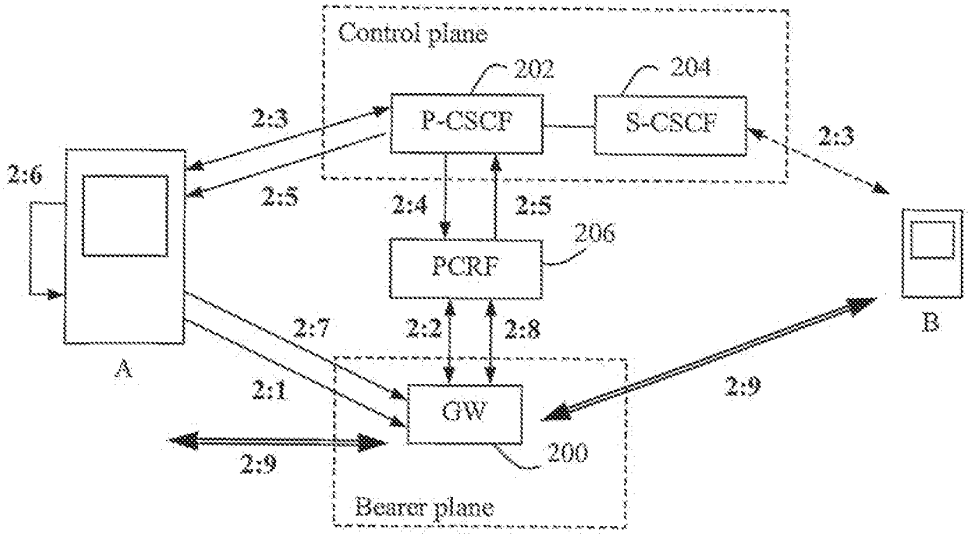

Control plane
202
204
P-CSCF
S-CSCF
2:3
2:3
2:6
2:5
2:4
2:5
206
PCRF
B
2:7
2:2
2:8
A
2:1
GW
2:9
2:9
200
Bearer plane

Fig. 2

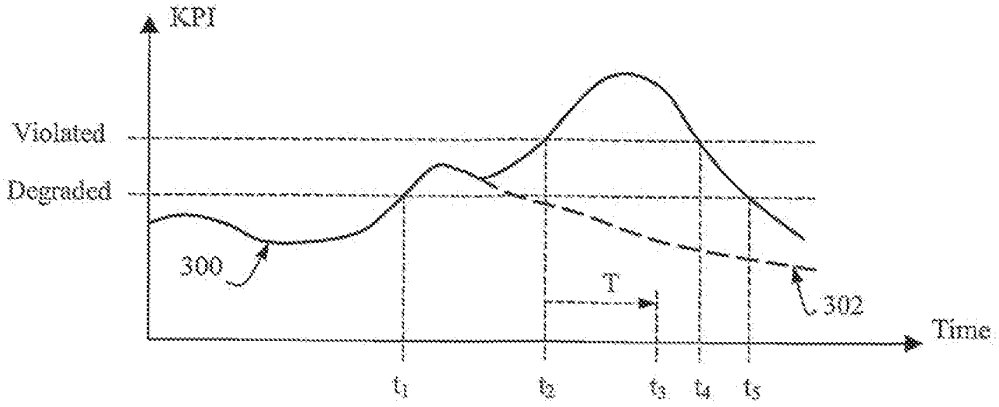

KPI
Violated
Degraded
300
T
302
Time
t1
t2
t3
t4
t5

Fig. 3

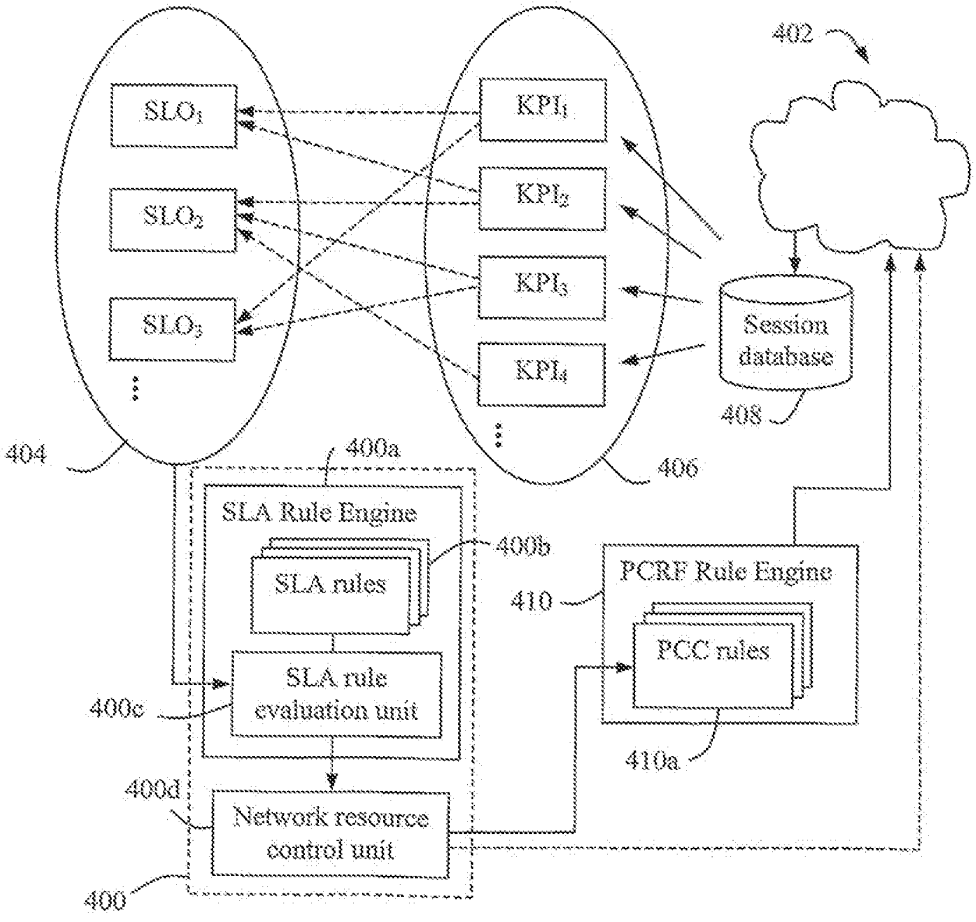

402
SLO1
SLO2
SLO3
KPI1
KPI2
KPI3
KPI4
Session database
408
404
406
400a
SLA Rule Engine
400b
SLA rules
SLA rule evaluation unit
400c
400d
Network resource control unit
400
PCRF Rule Engine
410
PCC rules
410a

Fig. 4

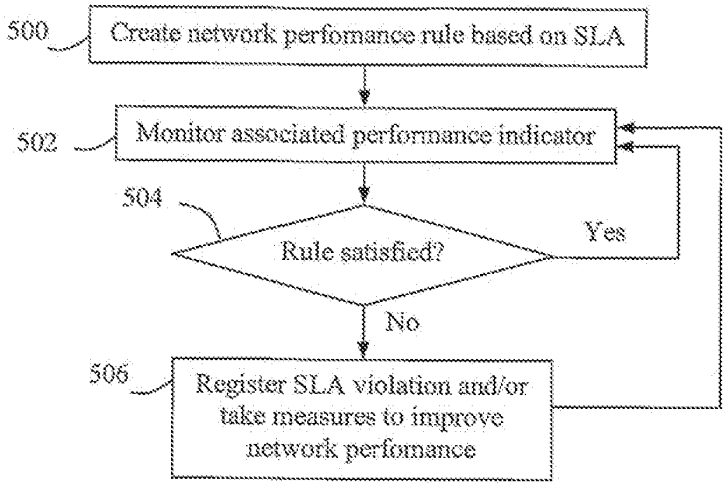

500
Create network performance rule based on SLA
502
Monitor associated performance indicator
504
Rule satisfied?
Yes
No
506
Register SLA violation and/or take measures to improve network performance

Fig. 5

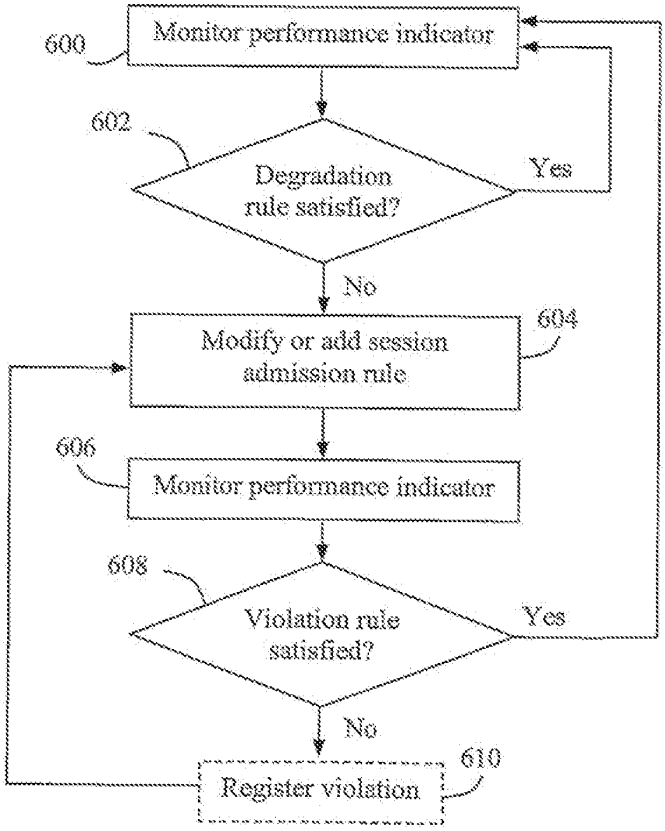

600
Monitor performance indicator
602
Degradation rule satisfied?
Yes
No
604
Modify or add session admission rule
606
Monitor performance indicator
608
Violation rule satisfied?
Yes
No
610
Register violation

Fig. 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE2006/001364

A. CLASSIFICATION OF SUBJECT MATTER

IPC: see extra sheet
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04Q, H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 0117169 A2 (ANDERSEN CONSULTING LLP), 8 March 2001 (08.03.2001), page 3, line 3 - line 14; page 38, line 9 - line 22; page 41, line 3 - line 28, page 47, line 2 - page 48, line 23; page 92, line 1 - page 95, line 11, claims 1-15, abstract	1-22
A	WO 03084133 A1 (NETWORK GENOMICS INC), 9 October 2003 (09.10.2003), paragraphs 0002-0004; 0011-0017, claims 1-6, abstract	1-22
A	EP 1259030 A2 (GATEWAY INC), 20 November 2002 (20.11.2002), paragraphs 0002-0009	1-22

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 July 2007	23 -07- 2007
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Elisabet Aselius/CM Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (April 2007)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE2006/001364

International patent classification (IPC)
H04Q 3/00 (2006.01)
H04L 12/24 (2006.01)

Download your patent documents at www.prv.se
The cited patent documents can be downloaded at www.prv.se by following the links:

- In English/Searches and advisory services/Cited documents (service in English) or
- e-tjänster/anförda dokument (service in Swedish).

Use the application number as username.
The password is **EBVIQUCPIL**.

Paper copies can be ordered at a cost of 50 SEK per copy from PRV InterPat (telephone number 08-782 28 85).

Cited literature, if any, will be enclosed in paper form.

WO	0117169	A2	08/03/2001	AU	7109200	A	26/03/2001
				TW	226164	B	01/01/2005
WO	03084133	A1	09/10/2003	AU	2003228411	A	13/10/2003
				US	20040153563	A	05/08/2004
EP	1259030	A2	20/11/2002	JP	2002374303	A	26/12/2002
				US	20020174217	A	21/11/2002

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 May 2006 (18.05.2006)

PCT

(10) International Publication Number
WO 2006/050758 A1

(51) International Patent Classification[7]: H04L 29/06

(21) International Application Number:
PCT/EP2005/002837

(22) International Filing Date: 17 March 2005 (17.03.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/626,691 10 November 2004 (10.11.2004) US

(71) Applicant *(for all designated States except US)*: TELEFONAKTIEBOLAGET LM ERICSSON (publ) [SE/SE]; S-164 83 Stockholm (SE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: BOMAN, Krister [SE/SE]; Nimbusgatan 3, S-431 44 Mölndal (SE).

(74) Agents: BERGENTALL, Annika et al.; Cegumark AB, P.O. Box 53047, S-400 14 Göteborg (SE).

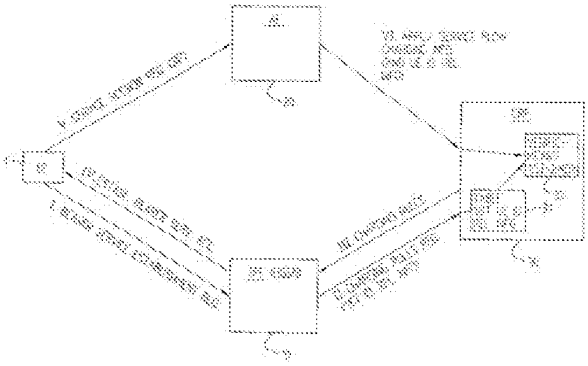
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: AN ARRANGEMENT, NODES AND A METHOD RELATING TO SERVICES ACCESS OVER A COMMUNICATION SYSTEM



(57) Abstract: The present invention relates to an arrangement in a communications system participating in user station (UE;1) service request and/or access procedures and comprising a number of packet data support nodes (10), a number of charging and/or policy handling nodes (30) and a number of application functions (20) handling mobility management and call control of mobile user stations requesting and/or accessing services. The packet data support node(s) (10) comprise(s) means adapted to send first mobile user station identity related information over a first interface (Gx,Gy;Gx/Gy) to a charging and/or policy handling node (30), at reception of a request for bearer services from a mobile user station (1). The application function(s) (20) comprise(s) means for, at reception of a request for a service session (SIP) from a mobile user station (1), sending second mobile user station identity related information to the charging and/or policy handling node (30), over a second interface (Rx,Rx/Gq). The policy and/or charging handling node (30) comprises verification means (32) adapted to establish whether the request for a bearer service to the packet data support node (10) and the request for a service session to the application function (20) (AF;P/S-CSCF) originate from one and the same mobile user station (1).

WO 2006/050758 A1

Title:

AN ARRANGEMENT, NODES AND A METHOD RELATING TO SERVICES ACCESS OVER A COMMUNICATION SYSTEM

FIELD OF THE INVENTION

The present invention relates to an arrangement in a communications system participating in user station service request or access procedures and it comprises a number of packet data support nodes, a number of charging and/or policy handling nodes and a number of application functions handling mobility management and call control of mobile user stations requesting or accessing services. The invention also relates to a node in a communications system supporting communication of packet data which is arranged to communicate with an application function and a packet data support node.

Still further the invention relates to a method involved in service request/access, in a communications system supporting communication of packet data and comprising a number of packet data nodes, a number of charging and/or policy handling nodes and a number of application functions handling mobility management and call control of mobile user stations requesting/accessing services.

STATE OF THE ART

So called IM (Internet IP Multimedia) services are getting more and more attractive and are already widely used. 3GPP IMS (IP Multimedia Subsystems) provides an IP-based session control

capability which is based on the SIP protocol (Session based Initiation Protocol). IMS may for example be used for the provisioning of services such as push-to-talk, instant messaging, presence and conferencing. IMS in UMTS (Universal Mobile Telecommunication System) will support IP multimedia applications such as video, audio and multimedia conferences. As referred to above, 3GPP will use the SIP, Session Initiation Protocol, as the signalling protocol for creating and terminating multimedia sessions, c.f. for example IETF RFC (Request For Comments) 3261, which deals with how the SIP signalling is protected between the subscriber and the IMS, how the subscriber is authenticated and how the subscriber authenticates the IMS.

The current security architecture for IMS is specified in 3GPP TS 33.203. However, it is a problem that implementations of the services referred to above will exist which are not fully compliant with 3GPP IMS. 3GPP IMS for example exclusively uses IPv6 (Internet Protocol version 6) although there may still be IMS implementations based on IPv4. However, the use of IPv4 instead of IPv6 is only one of the differences between "early" IMS implementations and implementations complying with 3GPP IMS to a full extent. It has been recognized that non-compliance with 3GPP TS 33.203 security features, will mainly produce problems at the user equipment (UE) due to the potential lack of support of the USIM/ISIM interface, particularly for user stations only complying with 2G. It has therefore been realized that, for such "early" IMS implementations, mechanisms are needed to protect against the most significant security threats.

Such a temporary solution for "early" IMS implementations is suggested in 3GPP TR 33.878. The approach in this document is that the GGSN (Gateway GPRS Support Node) shall send MSISDN and

the IP address of a user station (User Equipment UE) towards a Radius server over the Gi interface. The Radius server then forwards the MSISDN and IP address to the HSS (Home Subscriber Server). When the user station, or the UE, sends a SIP register request towards a Serving-Call/Session Control Function (S-CSCF), including the IM (IP Multimedia) private identity, the S-CSCF will query the HSS, which, based on the IMS identity can identify the matching MSISDN and IP address of the UE and return the IP address to the S-CSCF. The S-CSCF will check with the stored IP address in the S-CSCF to see if it matches the received IP address from the UE. If the two IP addresses are the same, the S-CSCF can proceed with the registration procedure but if they are not the same, presumably due to an attack being in progress, the session is terminated. This gives a possibility to avoid certain attacks such as for example an attacker using the IMS identities of a victim, which may have as a consequence that the attacked user has to pay for the service accessed by the attacker.

An operator also runs the risk that the attacker does not pay for the bearer, e.g. a conversational bearer, due to the fact that for example FBC (Flow Based Charging) is implemented and zero rates the bearer. This means that for some use cases, an attacker could use certain services for free.

3GPP TR 33.878 assumes that "During a PDP-context request towards the IMS, the GGSN shall send a "Radius Accounting-Request-Start" message to a Radius server attached to the HSS. This may lead to the assumption that an operator uses a specific IMS APN (Access Point Name) or that the UE uses a signalling PDP-context. It is however believed that the assumption relating to the use of a specific IMS APN is too restrictive from an architectural point of view. 3GPP TS 23.228 stipulates that

"When the UE uses GPRS-access for IMS services, it shall be able to establish a dedicated signalling PDP-context for IM subsystem related signalling or utilize a general-purpose PDP-context for IM subsystem signalling traffic". There is hence no guarantee that the UE uses a signalling PDP-context for IMS, may lead to the conclusion that the GGSN should then send the IP address of a UE towards a Radius server for all services. Further, an UE may have several IP addresses towards one and the same GGSN and it appears necessary that the GGSN, in such a case, has to send all these IP addresses towards the Radius server. Furthermore, in the general case a UE may be attached to more than one GGSNs or it may have activated more than one IP network connections. Such scenarios are not discussed in 3GPP TR 33.878.

Furthermore, the solution discussed in this document suggests the introduction of a so called "idle timer" in GGSN. This has an impact on how IP addresses towards an UE are allocated or released and an IP address can be reused for a user when re-establishing a PDP-context at a later stage. Apparently the reason for the introduction of the idle timer is to reduce the load on the HSS and, when there is no PDP-context available between GGSN and an UE, the GGSN stores or reserves the allocated IP address for a specific UE for a time which is equal to the time period set in the idle timer. This means that the GGSN does not upon termination of all bearers send an accounting-request stop until the idle timer has expired. Thus the idle timer has not been introduced for security reasons but in order to reduce the communication towards the HSS which is responsible for the creation of authentication vectors and hence the intention is to reduce communication with the HSS, or not to allow more communication with HSS than necessary in order to avoid that it becomes overloaded, which might have the effect that other services, for example voice, would suffer. The use of

the idle timer is disadvantageous for example in that there will be an increasing risk for conflicts between dynamically allocated IP addresses. Furthermore resources are held more than necessary.

The suggested solution is also disadvantageous in that more and more information is sent from GGSN over the Gi interface which will become more and more loaded. It is also disadvantageous that a Radius server is needed in order to verify the IP addresses and, as referred to above, it does not take general architectures into account such as when a UE for example is attached to more than one GGSN and/or has activated more than one IP network connection etc.

SUMMARY OF THE INVENTION

Therefore an arrangement as initially referred to is needed through which IMS security can be provided for in a straight-forward and easy manner. Particularly an arrangement is needed through which so called "early" IMS security can be provided, i.e. in networks still not (fully) complying with the 3GPP TS 33.203 Release 5. Particularly an arrangement is needed through which attacks can be detected and prevented and particularly through which security can be increased for end users as well as for operators, for example such that an attacker is prevented from using certain services for free or making someone else pay for it, and that an attack can be detected as early as possible which is extremely important for the end user as well as for the operator.

Particularly an arrangement is needed which is flexible and which can be adapted to different network environments, such as for example so called aware network environments. Still further an arrangement is needed which takes into account and takes

advantage of the development for example within 3GPP or similar systems. Still further an arrangement is needed through which the amount of information sent over the so called Gi interface from Gateway GPRS Support Nodes or similar can be reduced instead of increased. Most particularly an arrangement is needed through which there is no need for the introduction of additional nodes and more generally through which the number of nodes involved and affected can be kept as low as possible. An arrangement is also needed which reduces or even eliminates the drawbacks and impacts on IP address allocation or release towards an UE arising through the introduction of a so called idle timer in a GGSN.

Still further an arrangement is needed which can be used in a general architecture and which takes into account situations, as for example when a user equipment is attached to more than one GGSN (or CGSNs or other gateways) and/or when a user equipment has activated more than one IP network connection. Still further a node is needed which assists in provisioning of one or more of the objects referred to above. Still further a method is needed through which one or more of the abovementioned objects can be achieved.

Therefore an arrangement as initially referred to is provided wherein the packet data support node(s)/comprise(s) means adapted to send first mobile user station identity related information over a first interface (Gx,Gy;Gx/Gy) to a charging and/or policy handling node, at reception of a request for bearer services from a mobile user station, the application function(s) comprise(s) means for, at reception of a request for a service session (SIP) from a mobile user station (UE), sending second mobile user station identity related information to the charging and/or policy handling node, over a second interface (Rx,Rx/Gq), and wherein the

policy and/or charging handling node comprises verification means adapted to establish whether the request for a bearer service to the packet data support node and the request for a service session to the application function (AF;P/S/I-CSCF) originate from one and the same mobile user station (UE). Particularly the first mobile user station identity related information comprises one or more of MSISDN, IMSI and IP address of the mobile user station, and the second mobile user station identity related information comprises IMS private identity (IMPI) and/or IMS public identity (IMPU), or IMSI and/or MSISDN, e.g. derived from IMS private or public identity (IMPI,IMPU), or the application function being adapted to fetch said information externally, e.g. from a HSS.

Particularly the IP addresses are received in, or derived from, the first mobile user station related information and the second mobile user station identity related information, and the verification means are adapted to compare the IP address obtained from the application function (AF) with the IP address obtained from the packet data support node. In one embodiment the verification means are adapted to, if required, establish and, use IMSI and/or MSISDN of said first and second mobile user station identity related information respectively, to compare the associated IP addresses. The charging and/or policy handling node may also be adapted to deduce IMSI and/or MSISDN from a received IMS private and/or public user identity (IMPI/IMPU).

In an advantageous embodiment the charging and/or policy handling node is adapted to construe a private user identity using IMSI of a mobile user station, e.g. when no ISIM application is implemented and further it is particularly adapted to identify and install charging rules to be applied at reception of a request therefor from a packet data support node.

According to the invention, if the first and second user identity related information do originate from one and the same mobile user station, the charging and/or policy handling node is particularly adapted to implement the provisioned charging rules. If, on the other hand, the first and second user identity related information do not originate from one and the same mobile user station, the charging and/or policy handling node may be adapted to reject the charging rules and/or deactivate implemented charging rules and discard concerned IP flow traffic, and further be adapted to inform the application function (AF) that no PDP-context is available.

Alternatively, if the first and second user identity related information do not originate from one and the same mobile user station, the charging and/or policy handling node may be adapted to select whether to or deactivate the charging rules in the packet data support node, or to keep them active (or activate them). In a particular embodiment the packet data support node comprises a node handling Traffic Plane Functions (TPF), e.g. a GGSN implementing TPF or a CGSN implementing TPF. It may alternatively comprise a stand-alone node or a gateway node handling TPF or TPF and policy enforcement (PEP).

According to different embodiments the charging and/or policy handling node comprises a CRF, a PDF or a PCCN. Particularly the first interface is Gx or Go/Gx merged with Gy if the charging and/or policy handling node is merged with an OCS, and the second interface comprises Rx or Rx merged with Gq.

A node (functional, logical or physical) in a communication system as initially referred to is also provided which is adapted to receive first mobile user station identity related information from a packet data support node relating to a bearer service

request for a mobile user station, or at reception of a request for a bearer service from a mobile user station (UE) in the packet data support node, to receive second mobile user station identity related information from an application function relating to a service session or at reception of a request for a service session therein, and comprises verification means for establishing whether the request for a bearer service (to the packet data support node) and the request for a service session (to the application function) originate from the same mobile user station. Particularly the first mobile user station identity related information comprises one or more of MSISDN, IMSI and IP address of the mobile user station. Even more particularly the second mobile user station identity related information comprises IMS private identity (IMPI) and/or IMS public identity (IMPU) of the mobile user station or IMSI and/or the handling node is adapted to derive IMSI and/or MSISDN from IMPI and/or IMPU and/or adapted to request said IMSI and/or MSISDN from the application function.

The verification means are preferably adapted to compare the IP address corresponding to or included in the first mobile user station identity related information and the IP address corresponding to or included in the second mobile user station identity related information, and particularly to, if required establish and, use IMSI and/or MSISDN of said first and second mobile user station identity related information, to compare the associated IP addresses. Alternatively it is adapted to derive IMSI and/or MSISDN from a received IMS private and/or public user identity (IMPI/IMPU). Particularly the charging and/or policy handling node is adapted to construe a private user identity using IMSI of a mobile user station, e.g. when no IMSI application is implemented.

It may also be adapted to identify and install charging rules to be applied at reception of a request therefor from a packet data support node. Particularly it is adapted to identify and install charging rules to be applied at reception of request therefor from a packet data support node, and if the first and second user identity related information do originate from one and the same mobile user station, to implement the provisioned charging rules, and, if the first and second user identity related information do not originate from one and the same mobile user station, to reject the charging rules and/or deactivate implemented charging rules and discard concerned IP flow traffic, to inform the application function (AF) that no PDP-context is available or, to select whether to deactivate the charging rules in the packet data support node, or keep them active, if activated. Particularly it comprises a CRF, a PDF or a PCCN or a node with a similar functionality.

Still further a method as initially referred to is provided which comprises the steps of; receiving a request for bearer services from a mobile user station in a packet data support node; sending, from the packet data support node, first mobile user station identity related information of the mobile user station to a charging and/or policy handling node; receiving, subsequently, before or substantially simultaneously, a request for a service session from the mobile user station in an application function; sending, from the application function, second mobile user station identity related information to the charging and/or policy handling node; establishing, in the charging and/or policy handling node, whether the request for a bearer service to the packet data support node and the request for a service session to the application function originate from the same mobile user station.

As referred to earlier, particularly the first mobile user station identity related information comprises one or more of the MSISDN, IMSI and IP address of the mobile user station, and the second mobile user station identity related information comprises IMS private identity (IMPI) and/or IMS public identity (IMPU), or the second user identity related information comprises IMSI and/or MSISDN, e.g. derived from IMS private or public identity (IMPI,IMPU), or the application function is adapted to fetch said information from a HSS. Particularly the method further comprises the steps of; establishing or finding IMSI and/or MSISDN of said first and second mobile user station identity related information, comparing at least part of said first and said second mobile user station identity related information with each other, and even more particularly of comparing the IP addresses of said first and second mobile user station identity related information and/or deriving IMSI and/or MSISDN from received IMS private and/or public user identities (IMPI/IMPU), and/or construing, in the charging and/or policy handling node, a private user identity using IMSI of the mobile user station, e.g. when no ISIM application is implemented.

Advantageously it also comprises the step of; identifying and installing applicable charging and/or policy rules at reception of a request from a packet data support node. Most particularly the method further comprises the steps of; if the first and second mobile user station identity related information do originate from one and the same mobile user station; implementing the provisioned applicable charging rules and/or policy rules in the charging and/or policy handling node, otherwise, discarding concerned IP flow traffic; or selecting whether to keep active or deactivate the charging rules. In preferred implementations the packet data support node comprises a node handling Traffic Plane Functions, and is a stand-alone node, or a GGSN or a CGSN or a gateway node

handling TPF and policy enforcement (PEP), and the charging and/or policy handling node comprises a CRF, a PDF or a PCCN or a node with a similar functionality.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will in the following be further described, in a non-limiting manner, and with reference to the accompanying drawings in which:

Fig. 1 shows a "state of the art" solution,

Fig. 2 is a schematical block diagram illustrating an arrangement according to the present invention,

Fig. 3 is a simplified sequence diagram illustrating the provisioning of the first mobile user station identity related information to a charging and/or policy handling node,

Fig. 4 is a simplified sequence diagram including the procedure when the second mobile user station identity related information is provided to the charging and/or policy handling node,

Fig. 5 is a sequence diagram illustrating an exemplary procedure of the provisioning of the second mobile user station identity related information and verification thereof,

Fig. 6A is a sequence diagram describing an exemplary implementation and wherein the verification result is positive,

Fig. 6B is a sequence diagram describing the exemplary implementation of Fig. 6A but wherein the verification result is negative,

Fig. 7 is a schematical block diagram illustrating the inventive concept implemented in one general architecture,

Fig. 8A illustrates another general architecture in which the inventive concept can be implemented,

Fig. 8B shows still another architecture in which the inventive concept can be implemented, and

Fig. 9 is a flow diagram describing the procedure according to one embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Fig. 1 shows the state of the art solution discussed earlier in the application for providing security for IMS accesses, particularly for early implementations of IMS not yet complying with the requirements of 3GPP TS 33.203. Such a solution is among others discussed in 3GPP TR 33.878. It is supposed that an UE l_0 requests a bearer from GGSN 10_0. PDP-contexts are set-up in a conventional manner, primary or secondary depending on which is the relevant procedure. The user equipment UE l_0 may have several IP addresses, one IP address per IP network connection. GGSN comprises an idle timer 11_0 through use of which an IP address can be reused for a user when re-establishing a PDP-context at a later stage. There is no specific IMS APN (Access Point Name) and GGSN cannot assume use of a signalling PDP-context. Knowing, in a conventional manner MSISDN and the IP address of UE l_0, GGSN 10_0 sends MSISDN and the IP addresses over

the Gi interface towards Radius Server 30_0, i.e. all UE IP addresses and MSISDN are sent. The idle timer used in GGSN 10_0 apparently is intended mainly to reduce the load on the HSS. This is suggested to be in the order of hours working in such a way that when there is no PDP-context available between GGSN and UE, the GGSN stores or reserves the allocated IP address for a specific UE for a time period which corresponds to the time set in the idle timer. This would mean that the GGSN does not upon termination of all bearers send any accounting-request-stop until the idle timer has actually expired. This idle timer does influence how IP addresses towards the UE are allocated or released.

The Radius Server 30_0 forwards the information, i.e. MSISDN and the IP address to HSS 40_0. All UE IP addresses and the MSISDN are sent according to the idle timer 11_0 in the GGSN 10_0, i.e. the sending will generally be slowed down, or delayed, until the idle timer expires.

Then a request for a SIP registration (SIP REGISTER) is sent from UE 1_0 to S-CSCF, the S-CSCF sends the IMS identity onto the HSS 40_0. HSS 40_0 returns a reply to S-CSCF with the registered IP address back to the S-CSCF 20_0. S-CSCF 20_0 comprises a verification means 21_0 or a functionality for performing a verification through verifying that the IP address from HSS 40_0 is the same as the stored IP address in S-CSCF. If they are not the same, the session is terminated because it is assumed that an attack is ongoing.

The disadvantages of such a solution has been discussed earlier in the application.

Fig. 2 shows an arrangement according to one implementation of the inventive concept according which a FBC (Flow Based Charging) architecture is reused. The problem that is solved through the arrangement consists in the provisioning of a mechanism ensuring that the network is enabled to verify that the IP address of a user equipment UE 1 and the IMS identity of the UE 1 at SIP layer correspond to the same user at the bearer level. According to invention, the involved nodes or functions are the GGSN 10 handling traffic plane function TPF, (according to the inventive concept it is the traffic plane function that is used and it may be realized as a stand-alone node, provided in a GGSN or in a CGSN or in any node at least including the TPF functionality), a charging and/or a policy handling node 30 which in this embodiment is implemented as a CRF (Charging Rules Functionality) and an application function AF 20 which may be implemented as a S-CSCF, a P-CSCF or similar, and the concerned user equipment UE 1. CRF 30 comprises storing means, e.g. a table for storing first user equipment identity related information as will be explained below, and verification means 32 for verifying if at least a given part of a first UE identity related information corresponds to second UE identity related information, e.g. through comparing at least parts of said first and second identity related information respectively. By at least part of is meant that it might not be necessary to compare all the information depending on implementations, as also will be explained below. It is, however, in Fig. 2 supposed that UE 1 sends a request for bearer service establishment (I) to TPF 10 which may be implemented in a GGSN. TPF 10 then, in for example a CRF based implementation, as in Fig. 2, requests for charging rules (II) from CRF 30. First UE identity related information accompanies the request. The first UE identity related information may be provided from the UE or provided from TPF 10 (if it for example relates to a secondary PDP-context request).

How this is done or acquired is not of importance for the functioning of the present invention, the main thing being that the first UE identity related information is available in TPF 10. The first UE identity related information accompanying the charging rules request is then stored into table 31 in CRF 30, CRF 30 among others using the first UE identity related information and possibly other information, for example QoS information about application function etc., identifies which are the relevant charging rules to be installed, and then provides the charging rules to TPF 10 (III). TPF 10 performs the relevant charging rule actions or provides for implementation of the provisioned charging rules, and sends an acceptance message relating to the establishment of the bearer service request, (IV) to UE 1.

Once a bearer has been established, UE 1 sends a request for a service session using SIP, (V) to AF 20. In GPRS systems, and from an FBC point of view, a proxy-CFCS and/or a serving-CFCS and/or an Interogating CSCF may, depending on operator configuration, act as an AF operating the Rx interface and/or the Gq interface. In the request using the SIP protocol, second user identity related information of the requesting is included. AF 20 then sends application/service flow charging information, including the second UE identity related information to CRF 30 (VI). If for example an S-CSCF acts as an AF, it may, before a registration session is proceeded, unsolicited forward the private user identity along with the IP address of the UE over Rx towards CRF 30. At PDP-context activation, FBC assumes that the IP address of the UE and IMSI (or MSISDN) are sent towards the CRF 30 over the Gx interface. Based on the above assumptions, CRF 30 is then aware of the IMSI both from the AF 20, particularly S-CSCF, and the TPF 10, particularly GGSN, and the corresponding IP addresses and hence capable to, in the

verification means 32, check if the IP addressee match or originate from one and the same user (in Fig. 1 they actually both originate from the same UE, namely UE 1). If they do not correspond, or do not originate from one and the same end user, i.e. if there is no match between the IP addresses, CRF 30 (in this case; it should be clear that for policy handling the node could be a PDF, but it may also be a PCCN for handling policy and/or charging) notifies the AF 20, particularly S-CSCF, in order to provide for a termination of the SIP session, and the CRF 30 should deactivate any corresponding rule(s) at the TPF 10.

It could of course also be optional whether the session should be terminated or not and optionally it may be operator controlled, such that under certain circumstances for example an operator might want to localize or track an attacker and therefore would let the session proceed. Also for other reasons an operator might allow the session under certain conditions etc. The invention is not limited to any particular measures to be taken, but it is fundamental that an awareness is provided concerning the correspondence of addresses and possible attacks or similar.

The second user identity related information sent from AF 20 to CRF 30 may for example contain IMPI (IMS Private Identity), IMPU (IMS Public Identity), and the IP address. Alternatively the AF 20 may send IMSI or MSISDN directly. This can either be derived from the private identity or alternatively the AF, e.g. S-CSCF may contact HSS to identify IMSI or MSISDN.

If for example the UE does not contain an ISIM, (but for example an USIM) IMPI may be derived and the private user identity

should take the form of a NAI and shall have the form username@realm as specified in IETF RFC 2486, clause 3.

A representation of the IMSI may be contained within the NAI for the private identity.

If there is no ISIM application, the private user identity is not known and it may then be derived from IMSI. The private user identity may be construed out of the IMSI by using the whole string of digits as username part of the private user identity, and by converting the leading digits of the IMSI, i.e. MNC and MCC into a domain name.

The result will be a private user identity of the form: "<IMSI>@ims.mnc<MNC>.mcc<MCC>.3gppnetwork.org". For example: If the IMSI is 234150999999999 (MCC=234, MNC=15), the private user identity will takes the form 234150999999999@ims.mnc015.mcc234.3gppnetwork.org

Hence, CRF/PDF (PCCN) is capable to, based upon the reception of IMPI, i.e. the IMS Private Identity, or even IMPU, the Public User Identity, and IP address from an AF, for example a CSCF such as P-CSCF, S-CSCF, or I-CSCF, over Rx or Gq or Rx merged with Gq, derive IMSI (or MSISDN) of the UE. It should be noted that the AF alternatively may send MSISDN and/or IMSI directly over Rx/Gq as referred to above.

Furthermore, since the TPF, here TPF implemented in a GGSN, shall have the capability to send the IMSI or MSISDN to the CRF/PDF/PPCN (over Gx or Gx/Gy if merged), the CRF/PDF can then match the IMSI (or MSISDN) from CRF and TPF (here GGSN) with the respective IP addresses, i.e., CRF 30 is capable to find IMSI (MSISDN) from, here, GGSN and AF and check that it is the same

IP address associated. If there is a match, the CRF preferably implements, in this case, charging rules as specified in 3GPP TS 23.125. If, on the other hand, the IP addresses do not originate from one and the same UE, the CRF may decide not to implement the charging rules or deactivate already implemented charging rules as referred to above, and hence the traffic may be discarded for the concerned IP flow. The CRF then informs the AF, e.g. a P-CSCF/S-CSCF that no PDP context is available.

Alternatively, may, in certain situations, as also referred to above, an operator want to allow the traffic, for example in order to track an attacker. Hence it may be an option for the charging and/or policy handling node 30 to keep rules in the TPF 10 active or to deactivate them.

In some implementations the TPF is implemented in a GGSN (or CGSN) but, as referred to above, TPF may be implemented in a stand-alone node or function or be allocated elsewhere than in a GGSN/CGSN, the main thing being that it is the traffic plane function TPF that is involved according to the inventive concept, since it is TPF that operates Gx towards CRF/PDF (or PCCN) (or Gx/Gy when merged).

The invention functions in a similar way for public user identities as for private user identities. 3GPP TS 23.003 specifies that when no ISIM is used, the temporary public user identity should be of the form "user@domain", and should therefore be equal to the private user identity. It is further possible to use a public user identity rather than the private identity at SIP layer. Then the AF, e.g. S-CSCF, may e.g. request an IMSI from HSS based on the public user identity, derive IMSI from the public user identity or send the public user identity towards CRF/PDF/PCCN that derives IMSI or sends a

query to a Subscription Profile Repository, if implemented, cf. Fig. 5, for getting IMSI or MSISDN.

Fig. 3 is a schematical sequence diagram illustrating the messaging between an UE, a TPF and, in this case a CRF. It should be clear, as referred to above, that the TPF can be provided for as a stand-alone node, in a GGSN, in a CGSN or in any other node or functioning means at least providing the TPF functionality. It should also be clear that instead of a CRF handling charging rules, it could as well be a PDF handling policy rules or a combined node or functionality handling e.g. both policy and charging. Hence, it is supposed that the UE, in a conventional manner, sends a request for establishment of a bearer service, 1, to e.g. a GGSN providing the TPF functionality. At PDP-context activation, first or primary, the GGSN, in this case, will send terminal identities such as one or more of IP address, MSISDN and IMSI towards CRF over the Gx interface in a request for charging rules, 2, i.e. the first user equipment identity related information is accompanying the request for charging rules. CRF stores the received first UE identity related information, 3, and identifies which are the charging rules that are applicable and should be installed, 4. Subsequently the charging rules are provisioned to TPF, 5. TPF at reception of the charging rules performs the charging rule action or actions, 6, and sends an establish bearer service acceptance message to UE, 7.

It should be clear that, as referred to above, the CRF may be merged with a PDF. Even more particularly CRF (and/or PDF or PCCN) may be merged with an Online Charging System OCS in which case Gx may be merged with Gy. These interfaces are defined in TS 23.125, and may e.g. in one embodiment both be based on Diameter Credit Control.

Fig. 4 is schematical sequence diagram illustrating the procedure through which the second UE identity related information is provided to CRF (for example) such that a verification can be carried out according to the inventive concept.

Hence, it is supposed that when UE sends a request for a service session using the SIP protocol, 8, to AF which e.g. may comprise a P-CSCF or an S-CSCF or an I-CSCF acting as an AF, this includes sending user identities to the CRF over Rx, which is described in 3GPP TS 23.228, which herewith is incorporated herein by reference. In other words, the UE sets up a SIP Register towards the IMS network. Hence, when a request for service session, SIP Register, 8, is sent to AF, AF provides application/service data flow charging information including second UE identity related information, e.g. IMPI, IMPU, IP address, to CRF, 9. CRF, having stored the first UE identity related information in storing means, fetches said first UE identity related information and compares it with the second UE identity related information received from the AF, 10. It is here supposed that the first and the second UE identity related information respectively originates from one and the same UE, and hence an acknowledgment, 11, is sent to AF and the service is setup implementing the relevant charging rules or similar.

It should be clear that the Rx interface may be merged with Gq and the AF may hence operate Rx/Gq. Rx and Gq may in particular embodiments both be based on Diameter Nasreq.

It should be clear that, as an alternative of sending IMPI and IMPU i.e. IMS private and public identity respectively as described in TS 23.228 and TS 23.003, the AF, or any CSCF, may

send IMSI or MSISDN directly to CRF. IMSI or MSISDN may be derived from the private identity or communication with HSS may be needed for identification of IMSI or MSISDN. This has been discussed earlier with reference to Fig. 2.

The CRF particularly comprises a table for translation between IMPI and IMSI.

Fig. 5 is a sequence diagram illustrating the procedure when the second UE identity related information is provided to e.g. a CRF or a PCCN etc., where (here) a successful verification is performed. It is here supposed that a PDP-context is established, and that the first UE identity related information is available in a CRF, PCCN etc. It is supposed that the UE is in a visited network and sends a SIP register, 80, with public user ID, private user ID, UE IP address to P-CSCF. The SIP register is sent from P-CSCF to I-CSCF of the UE home network, 81. I-CSCF sends a Diameter (in this particular embodiment) Cx-Query/Cx-Select-Pull message, 82, to HSS, which returns the corresponding response, 83, concerning IMS identities. SIP register with Public User ID, Private User ID, UE IP Address is then sent from I-CSCF to S-CSCF, 84. In this particular embodiment it is supposed that S-CSCF act as an AF. In alternative embodiments P-CSCF or I-CSCF may acts as AF. It shall be clear that here TPF and the subscription information repository are not visible.

Steps 85,86 relate to provisioning of security parameters, application IP address etc. which however is not of relevance for the inventive concept.

Subsequently it is supposed that the second UE identity related information, Public User ID, Private User ID, UE IP Address are

provided to CRF (PCCN etc.), 87, and after verification (here supposed to be affirmative) with the first UE identity related information (not shown here), an acknowledgement is sent to S-CSCF, 88. S-CSCF sends a positive verification message, 2XX OK to I-CSCF, 89, which forwards it to P-CSCF, 90, which in turn forwards it to UE, 91.

Fig. 6A shows one example of an implementation in which it is supposed that the 1st and 2nd UE identity related information can be positively verified (i.e. there is correspondence) and in which the Diameter protocol is used.

Thus, first the UE sends (here) an Activate PDP Context Request to (here) an SGSN, 1_1. In a conventional manner SGSN sends a Create PDP Context Request to GGSN, 2_1, which here uses a Radius Client to obtain IP addresses of UEs.

In this embodiment it is supposed that the Diameter protocol is used (cf. 3GPP TS 29.210), and GGSN sends a request for rules, here a Diameter CCR (Credit Control Request) with IMSI, MSISDN IP Address, over Gx to CRF (or PCCN etc.), 3_1. CRF responds with a Diameter CCA (Credit Control Answer) with charging rules over Gx, 4_1, to GGSN, which sends a Create PDP Context response to SGSN, 5_1. SGSN sends an Activate PDP Context Accept to UE, 6_1, with the allocated PDP Address (IP Address). As described earlier, it is supposed that the UE sends a SIP register to GGSN, 7_1, which checks for IP address spoofing (i.e. checks that the UE has not changed, i.e. uses another, IP address than the one it has been given). When it is verified that the UE uses the "correct" IP address, a SIP register is sent to P-CSCF, 8_1, which sends it to I-CSCF, 9_1. I-CSCF then sends the SIP register to C-CSCF (here acting as AF), 10_1.

In this embodiment it is supposed that S-CSCF sends a Diameter Message (Nasreg application) AAR to CRF, 11_1, over Rx and containing the public user ID of the UE, and the private user ID and the IP address of the UE. CRF maps private and/or public user identity to MSISDN and/or IMSI to retrieve the associated IP address. CRF then performs the verification, i.e. checks whether the information received from GGSN (cf. step 3_1) is the same, and acknowledges that the verification was successful (in this case) over Rx to S-CSCF in an AAA answer, 12_1.

Finally a SIP 2XX OK message is sent from S-CSCF to UE, 13_1.

Fig. 6B is similar to Fig. 6A with the difference that the mapping in CRF is unsuccessful. Hence steps 1_1-11_1 are similar to steps 1_1-11_1 in Fig. 6A and therefore bear the same reference numerals.

Upon establishing in CRF that the mapping is unsuccessful, i.e. that the first and second UE identity related information is not the same, an error message is sent over Rx (Diameter) to S-CSCF, 12_2. Optionally CRF may then send a Diameter CCA message over Gx to GGSN, 13_2, that the relevant charging rules should be removed. GGSN then confirms the removal to CRF, 14_2, and (in any case, irrespectively of whether steps 13_2, 14_2 are implemented or not) an error message is sent from S-CSCF to UE, e.g. a SIP 4XX Forbidden, 15_2.

It should be clear that the inventive concept is not restricted to the use of Diameter. Radius may also be used or any other appropriate protocol. As referred to above, the flow will also be different if a CGGN is used, instead of a GGSN and a SGSN which however is not explicitly illustrated since the alterations should be obvious.

As also referred to above, it does not have to be S-CSCF that acts as an AF, and instead of a CRF it may e.g. be a PDF, PCCN etc.

Fig. 7 shows a particular embodiment in which the charging and/or policy handling node comprises a PCCN 30A (Policy and Charging Control Node). PCCN 30A communicates with service data flow based credit control 36A or more generally OCS. If for example PCCN 30A receives an IMS identity, e.g. IMPI from an AF (not shown) it may request IMSI from the subscription profile repository 35 over the Sp reference point (or particularly Sp-interface). More generally, if PCCN 30A receives an IMS identity, it may query subscription profile repository 35 in order to get a connection between the IMS identity (IMPI or IMPU) and the MSISDN or IMSI received from the TPF, e.g. in GGSN. This is advantageous in that it also covers cases when users utilize an ISIM which may have another structure for the IMPI, which for example is not based on IMSI or MSISDN. In still another embodiment (not shown) PCCN (or PDF 33A or CRF 34A in PCCN as an alternative to a PCCN), may query the service data flow based credit control 36A in order to get a connection between the IMS identity and the MSISDN/IMSI received from the TPF.

Fig. 8A schematically illustrates another architecture in which the inventive concept is applicable. Such an architecture is for example described in TR 23.803. The charging and/or policy handling node here comprises a PCCN 30B with a PDF 33B and a CRF 34B communicating with an AF 20B over a merged Rx/Gq interface. The packet data support node including TPF functionality 11 here comprises a gateway 10B which in addition to the TPF functionality 11 also includes a PEP (Policy Enforcement Point)

functionality, 12. Communication between PCCN 30B and gateway GW 10B takes place over the merged interface Go/Gx. PCCN 30B communicates with OCS (Online Charging System) 36B with a Camel SCP 37B and Service Data Flow Based Credit Control means 38B over the Ry interface. GW 10B communicates with OSC 36B over the Gy interface. GW 10B also communicates with a charging gateway function $15B_1$ and a charging collection function $15B_2$ over the Gz interface.

Fig. 8B shows still another embodiment similar to that of Fig. 8A but with the difference that OSC and PCCN are merged into one node 30C, communicating with GW 10C over a merged interface comprising interfaces Gy/Go/Gx merged with one another. In other aspects Fig. 8B is similar to Fig. 8A and OCS/PCCN 30C communicate over a merged interface Gq/Rx with AF 20C and GW 10C communicates over Gz with charging gateway function $15C_1$ and charging collection function $15C_2$.

It should be clear that these figures are merely illustrated in order to show some additional examples on architectures in which the inventive concept can be applied.

Hence, according to the invention, through only a slight addition to existing FBC architectures or similar early IMS security can be provided for. Operators using FBC for non-IMS based services migrating towards IMS are given the opportunity to simply upgrade FBC instead of HSS and Radius server, which is very advantageous. The inventive concept involves no implications on for example a GGSN architecture e.g. IP address allocation or release as it is the case with the proposal given in TR 33.878 and there is also no need to introduce an idle timer in GGSN. According to the invention identities from IMS

can be connected to e.g. GPRS based identities by using Rx and Gx interfaces.

Fig. 9 is a flow diagram describing one implementation of the inventive concept. It is here supposed that an UE requests TPF for bearer service establishment, 100. In the request, the UE includes first UE identity related information, or if it is a secondary PDP-context request, the first UE identity related information is already stored in the node handling TPF. This is however known in the art.

TPF then sends a request for charging and/or policy rules to CRF or PDF, or a PCCN, depending on the applicable implementation, including said first UE identity related information, 101. CRF/PDF (or PCCN) stores the first UE identity related information in storing means, e.g. a table, 102. Subsequently CRF/PDF (or PCCN) identifies and installs, using said first UE identity related information and possibly other information as referred to above, the relevant charging and/or policy rules, 103. Then CRF/PDF (or PCCN) provides the charging and/or policy rules to TPF, 104. A Bearer Service Establishment Acceptance message is then sent from TPF to UE, 105. Once a bearer service has been established, it is supposed that UE sends a SIP message or a SIP request, e.g. a SIP invite, to an Application Function AF, 106. It may then be checked whether the private user identity is known from the SIP message, 107. If not, it may be derived somehow either in AF or in communication with some other node or function as discussed above, 107A. Then, as well as if the private user identity actually was known from the SIP message or SIP request, the second UE identity related information is sent to CRF/PDF or PCCN, 108. In CRF/PDF or PCCN it is thus then examined whether the first UE identity related information and the second UE identity related information

originate from one and the same UE, 109. If not, it is checked whether the charging and/or policy handling rules are to be deactivated, 109A, if yes, they are deactivated and/or the concerned traffic is discarded, 109B. If however the option is selected that the rules are not to be deactivated in step 109A, as well as if the first and the second UE identity related information respectively do originate from one and the same UE, the charging and/or policy rules are implemented, 110.

It should be clear that the invention is not limited to the specifically illustrated embodiments but that it can be varied in a number of ways within the scope of the appended claims.

CLAIMS

1. An arrangement in a communications system participating in user station service request and/or access procedures and comprising a number of packet data support nodes (GGSN;CGSN;TPF;10;10B;10C), a number of charging and/or policy handling nodes (CRF;TPF;PCCN;30;30A;30B;30C) and a number of application functions (AF;(P),(S),(I)CSCF;20;20B;20C) handling mobility management and call control of mobile user stations (UE;1) requesting and/or accessing services,

c h a r a c t e r i z e d i n

that the packet data support node(s) (10;10B;10C) comprise(s) means adapted to send first mobile user station (UE;1) identity related information over a first interface (Gx,Gy;Gx/Gy) to a charging and/or policy handling node (CRF;PDF;PCCN;30;30A;30B;30C), at reception of a request for bearer services from a mobile user station (UE;1),

that the application function(s) (AF;(P),(S),(I)CSCF;20;20B;20C) comprise(s) means for, at reception of a request for a service session (SIP) from a mobile user station (UE;1), sending second mobile user station identity related information to the charging and/or policy handling node (CRF;PDF;PCCN;30;30A;30B;30C), over a second interface (Rx,Rx/Gq),

and in that the policy and/or charging handling node (CRF;PDF;PCCN;30;30A;30B;30C), comprises verification means (32;32A) adapted to establish whether the request for a bearer service to the packet data support node (10;10B;10C) and the request for a service session to the application function (AF;P/S/I-CSCF;20;20B;20C) originate from one and the same mobile user station (UE;1).

2. An arrangement according to claim 1,

c h a r a c t e r i z e d i n
that the first mobile user station identity related information comprises one or more of MSISDN, IMSI and IP address of the mobile user station (UE;1).

3. An arrangement according to claim 1 or 2,
c h a r a c t e r i z e d i n
that the second mobile user station identity related information comprises IMS private identity (IMPI) and/or IMS public identity (IMPU).

4. An arrangement according to claim 1 or 2,
c h a r a c t e r i z e d i n
that the second mobile user station identity related information comprises IMSI and/or MSISDN, e.g. derived from IMS private or public identity (IMPI,IMPU), or that the application function (20;20B;20C) is adapted to fetch said information externally, e.g. from a HSS.

5. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the IP addresses are received in, or derived from, the first mobile user station related information and the second mobile user station identity related information, and in that the verification means (32;32A) are adapted to compare the IP address obtained from the application function (20;20B;20C) with the IP address obtained from the packet data support node.

6. An arrangement according to claim 5,
c h a r a c t e r i z e d i n
that the verification means (32;32A) are adapted to, if required, establish and, use IMSI and/or MSISDN of said first and second

mobile user station identity related information respectively, to compare the associated IP addresses.

7. An arrangement according to claim 4,
c h a r a c t e r i z e d i n
that the charging and/or policy handling node (30;30A;30B;30C) is adapted to deduce IMSI and/or MSISDN from a received IMS private and/or public user identity (IMPI/IMPU).

8. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the charging and/or policy handling node (30;30A;30B;30C) is adapted to construe a private user identity using IMSI of a mobile user station, e.g. when no ISIM application is implemented.

9. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the charging and/or policy handling node (30;30A;30B;30C) is adapted to identify and install charging rules to be applied at reception of request therefor from a packet data support node (10;10B;$15B_1$,$15B_2$;10C.

10. An arrangement according to claim 9,
c h a r a c t e r i z e d i n
that if the first and second user identity related information do originate from one and the same mobile user station (UE;1), the charging and/or policy handling node (20;30A,36A;30B,36B;30C) is adapted to implement the provisioned charging rules.

11. An arrangement according to claim 9,
c h a r a c t e r i z e d i n
that if the first and second user identity related information do not originate from one and the same mobile user station, the

charging and/or policy handling node (20;30A,36A;30B,36B;30C) is adapted to reject the charging rules and/or deactivate implemented charging rules and discard concerned IP flow traffic, and in that it is adapted to inform the application function (20;20B;20C) that no PDP-context is available.

12. An arrangement according to claim 9,
c h a r a c t e r i z e d i n
that if the first and second user identity related information do not originate from one and the same mobile user station, the charging and/or policy handling node (30;30A;30B;30C) is adapted to select whether to activate or deactivate the charging rules in the packet data support node.

13. An arrangement according to any one of the preceding claims,
c h a r a c t e r i z e d i n
that the packet data support node (10;10B;10C) comprises a node handling Traffic Plane Functions (TPF).

14. A user according to claim 13,
c h a r a c t e r i z e d i n
that the packet data support node comprises a GGSN (10) implementing TPF.

15. An arrangement according to claim 13,
c h a r a c t e r i z e d i n
that the packet data support node comprises a CGSN implementing TPF.

16. An arrangement according to claim 13,
c h a r a c t e r i z e d i n
that it comprises a stand-alone node or a gateway node (10B;10C) handling TPF or TPF and policy enforcement (PEP).

17. An arrangement according to any one of the preceding claims, c h a r a c t e r i z e d i n that the charging and/or policy handling node (30;30A;30B;30C) comprises a CRF, a PDF or a PCCN.

18. An arrangement according to any one of the preceding claims, c h a r a c t e r i z e d i n that the first interface is Gx or Go/Gx merged with Gy if the charging and/or policy handling node (30A;30B) is merged with an OCS (36A;36B).

19. An arrangement according to any one of the preceding claims, c h a r a c t e r i z e d i n that the second interface comprises Rx or Rx merged with Gq.

20. A charging and/or policy handling node (30;30A;30B;30C) in a communications system supporting communication of packet data and arranged to communicate with an application function (20;20B;20C) and a packet data support node (10;10B;10C), c h a r a c t e r i z e d i n that it is adapted to receive first mobile user station identity related information from the packet data support node (10;10B;10C) at reception of a request for a bearer service from a mobile user station (UE;1) in the packet data support node (10;10B;10C), to receive second mobile user station identity related information from the application function at reception of a request for a service session therein, and in that it comprises verification means (32;32A) for establishing whether the request for a bearer service to the packet data support node (10;10B;10C) and the request for a service session to the application function (20;20B;20C) originate from the same mobile user station (UE;1).

21. A charging and/or policy handling node according to claim 20, characterized in that the first mobile user station identity related information comprises one or more of MSISDN, IMSI and IP address of the mobile user station.

22. A charging and/or policy handling node according to claim 20 or 21, characterized in that the second mobile user station identity related information comprises IMS private identity (IMPI) and/or IMS public identity (IMPU) of the mobile user station or IMSI and/or that the handling node is adapted to derive IMSI and/or MSISDN from IMPI and/or IMPU and/or adapted to request said IMSI and/or MSISDN from the application function.

23. A charging and/or policy handling node according to any one of claims 20-22, characterized in that the verification means (32;32A) are adapted to compare the IP address corresponding to or included in the first mobile user station identity related information and the IP address corresponding to or included in the second mobile user station identity related information.

24. A charging and/or policy handling node according to claim 23, characterized in that the verification means (32;32A) are adapted to, if required establish and, use IMSI and/or MSISDN of said first and second mobile user station identity related information, to compare the associated IP addresses.

25. A charging and/or policy handling node according to claim 22,

c h a r a c t e r i z e d i n
that it is adapted to derive IMSI and/or MSISDN from a received IMS private and/or public user identity (IMPI/IMPU).

26. A charging and/or policy handling node according to any one of claims 20-25,
c h a r a c t e r i z e d i n
that it is adapted to construe a private user identity using IMSI of a mobile user station, e.g. when no IMSI application is implemented.

27. A charging and/or policy handling node according to any one of claims 20-26,
c h a r a c t e r i z e d i n
that it is adapted to identify and install charging rules to be applied at reception of a request therefor from a packet data support node (10;10B;10C).

28. A charging and/or policy handling node according to any one of claims 20-27,
c h a r a c t e r i z e d i n
that it is adapted to identify and install charging rules to be applied at reception of request therefor from a packet data support node (10;10B;10C), and if the first and second user identity related information do originate from one and the same mobile user station (1), it is adapted to implement the provisioned charging rules, and, if the first and second user identity related information do not originate from one and the same mobile user station,
it is adapted to reject the charging rules and/or deactivate implemented charging rules and discard concerned IP flow traffic, to inform the application function (20;20B;20C) that no PDP-context is available or, to select whether to deactivate the

charging rules in the packet data support node (10;108;10C) or to keep them active.

29. A charging and/or policy handling node according to any one of claims 20-28,
characterized in
that it comprises a CRF, a PDF or a PCCN or a node with a similar functionality.

30. A method in controlling service access, in a communication system supporting communication of packet data and comprising a number of packet data nodes, a number of charging and/or policy handling nodes and a number of application functions handling mobility management and call control of mobile user stations requesting/accessing services, comprising the steps of:

- receiving a request for bearer services from a mobile user station in a packet data support node,
- sending, from the packet data support node, first mobile user station identity related information of the mobile user station to a charging and/or policy handling node,
- receiving, subsequently, before or substantially simultaneously, a request for a service session from the mobile user station in an application function,
- sending, from the application function, second mobile user station identity related information to the charging and/or policy handling node,
- establishing, in the charging and/or policy handling node, whether the request for a bearer service to the packet data support node and the request for a service session to the application function originate from the same mobile user station.

31. A method according to claim 30,

wherein the first mobile user station identity related information comprises one or more of the MSISDN, IMSI and IP address of the mobile user station, and wherein the second mobile user station identity related information comprises IMS private identity (IMPI) and/or IMS public identity (IMPU), or the second user identity related information comprises IMSI and/or MSISDN, e.g. derived from IMS private or public identity (IMPI,IMPU), or the application function is adapted to fetch said information from a HSS.

32. A method according to claim 30 or 31,
further comprising the steps of:

- establishing or finding IMSI and/or MSISDN of said first and second mobile user station identity related information,
- comparing at least part of said first and said second mobile user station identity related information with each other.

33. A method according to claim 30,31 or 32,
comprising the step of:

- comparing the IP addresses of said first and second mobile user station identity related information.

34. A method according to any one of claims 30-33,
comprising the step of:

- deriving IMSI and/or MSISDN from received IMS private and/or public user identities (IMPI/IMPU).

35. A method according to any one of claims 30-34,
comprising the step of:

- construing, in the charging and/or policy handling node, a private user identity using IMSI of the mobile user station, e.g. when no ISIM application is implemented.

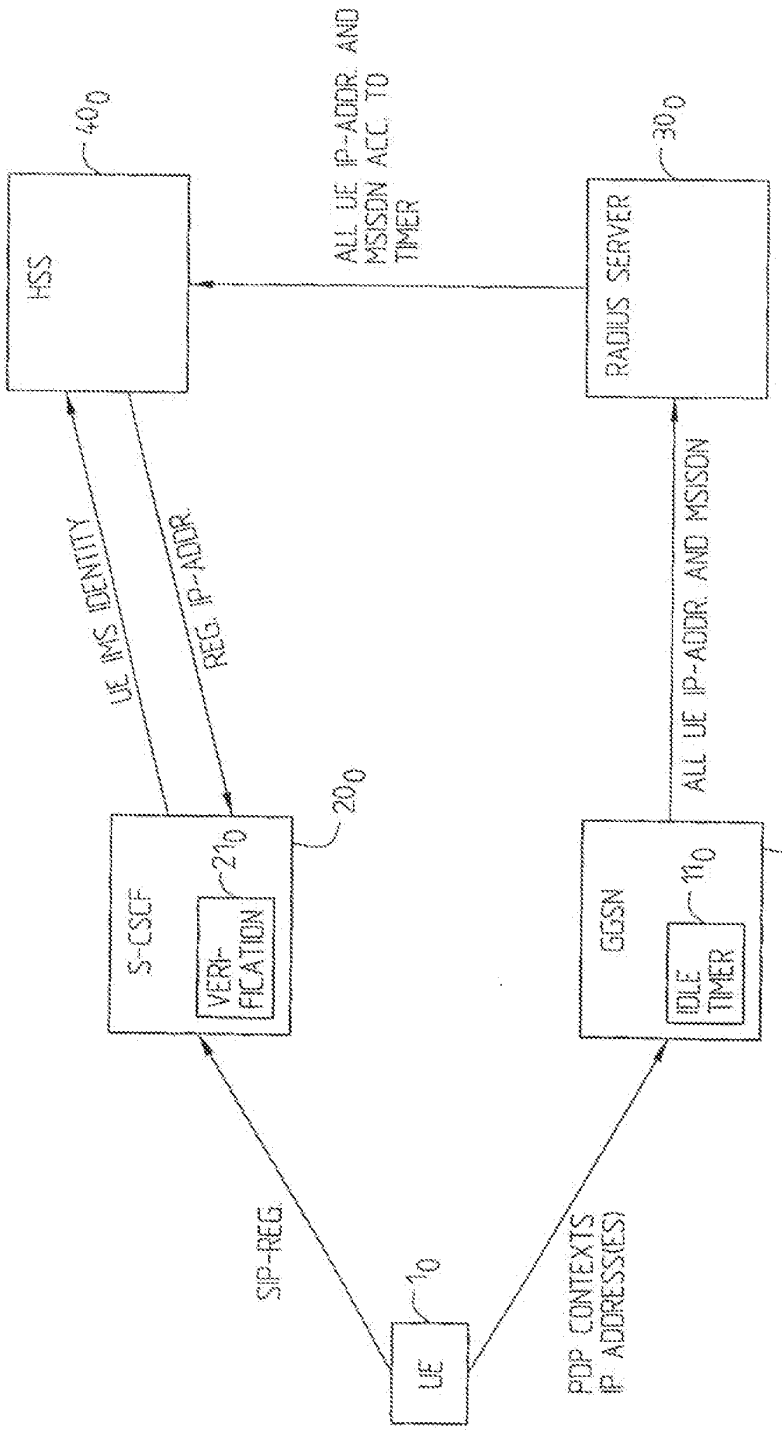
36. A method according to any one of claims 30-35, comprising the step of:

- identifying and installing applicable charging and/or policy rules at reception of a request from a packet data support node.

37. A method according to claim 36, further comprises the steps of:

- if the first and second mobile user station identity related information do originate from one and the same mobile user station,
- implementing the provisioned applicable charging rules and/or policy rules in the charging and/or policy handling node, otherwise,
- discarding concerned IP flow traffic, or
- selecting whether to activate or deactivate the charging rules.

38. A method according to any one of claims 30-37, wherein the packet data support node comprises a node handling Traffic Plane Functions, and is a stand-alone node, or a GGSN or a CGSN or a gateway node handling TPF and policy enforcement (PEP), and wherein the charging and/or policy handling node comprises a CRF, a PDF or a PCCN or a node with a similar functionality.



STATE OF THE ART

Fig. 1

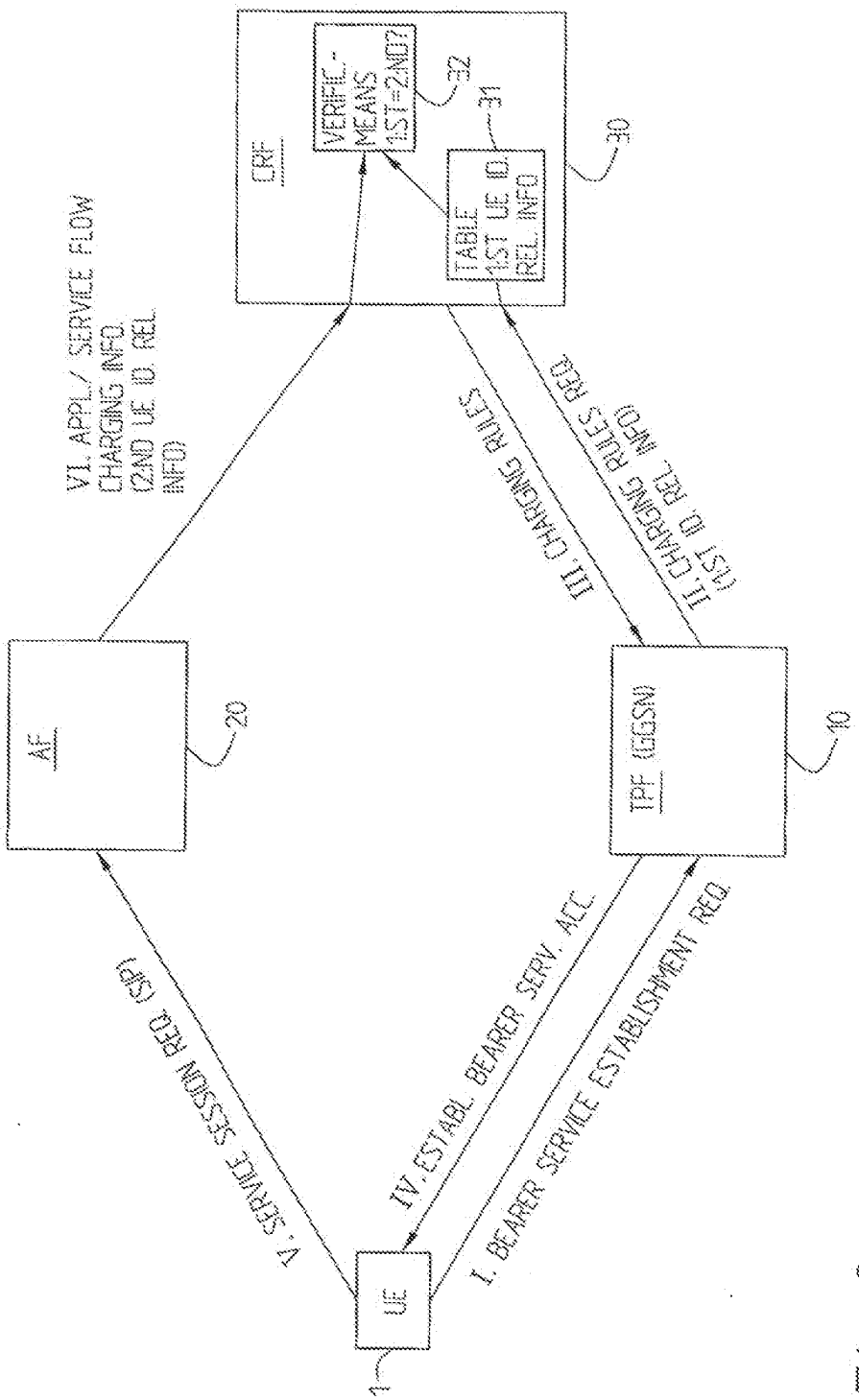

CRF
VERIFIC.-MEANS (1ST=2ND)
32
TABLE (1ST UE ID, REL INFO)
31
30
VI. APPL./ SERVICE FLOW CHARGING INFO. (2ND UE ID, REL INFO)
III. CHARGING RULES
II. CHARGING RULES REQ. (1ST ID, REL INFO)
AF
20
TPF (GGSN)
10
V. SERVICE SESSION REQ. (SP)
IV. ESTABL. BEARER SERV. ACC.
I. BEARER SERVICE ESTABLISHMENT REQ.
UE
1

Fig. 2

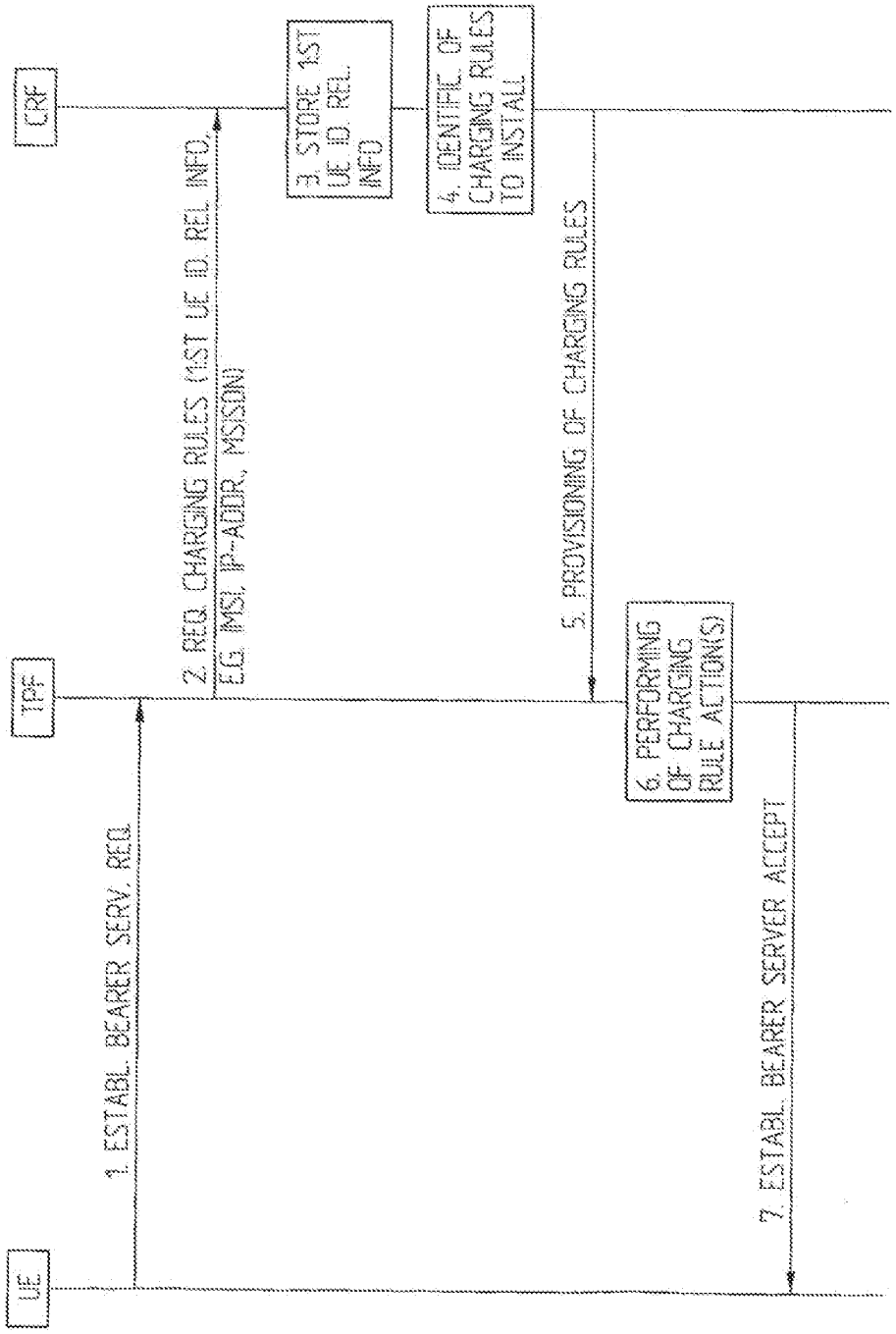

UE
TPF
CRF
1. ESTABL. BEARER SERV. REQ.
2. REQ. CHARGING RULES (1ST UE ID, REL INFO, E.G. MSI, IP-ADDR, MSISDN)
3. STORE 1ST UE ID, REL INFO
4. IDENTIFC. OF CHARGING RULES TO INSTALL
5. PROVISIONING OF CHARGING RULES
6. PERFORMING OF CHARGING RULE ACTION(S)
7. ESTABL. BEARER SERVER ACCEPT

Fig. 3

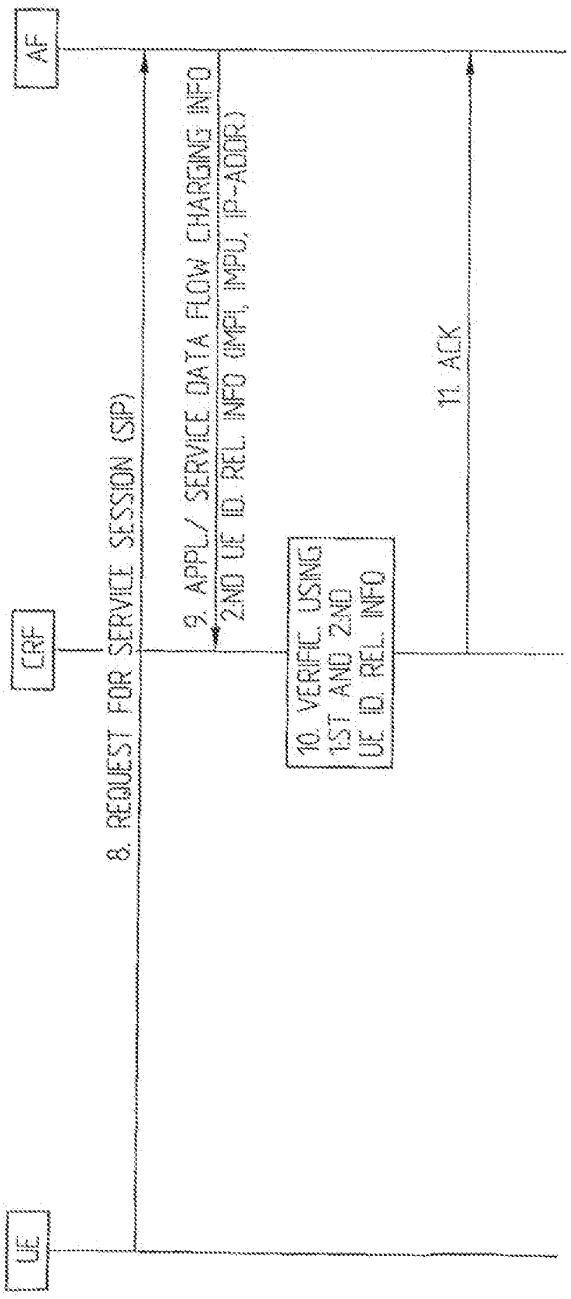

AF
CRF
UE
8. REQUEST FOR SERVICE SESSION (SIP)
9. APPL./ SERVICE DATA FLOW CHARGING INFO
2ND UE ID REL INFO (IMPI, IMPU, IP-ADDR)
10. VERIFIC. USING
1ST AND 2ND
UE ID REL INFO
11. ACK

Fig. 4

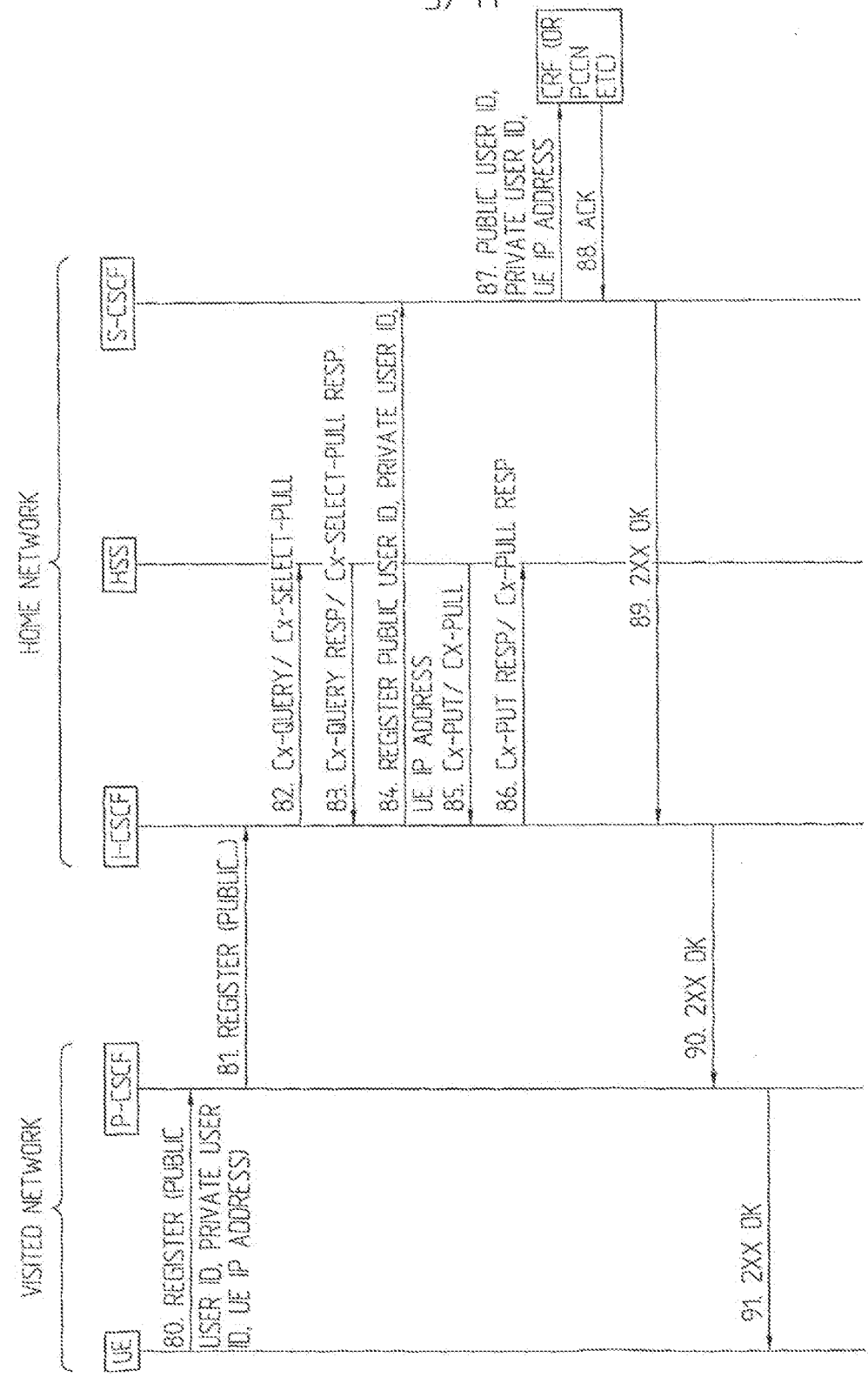

VISITED NETWORK
HOME NETWORK
UE
P-CSCF
I-CSCF
HSS
S-CSCF
80. REGISTER (PUBLIC USER ID, PRIVATE USER ID, UE IP ADDRESS)
81. REGISTER (PUBLIC...)
82. Cx-QUERY/ Cx-SELECT-PULL
83. Cx-QUERY RESP/ Cx-SELECT-PULL RESP.
84. REGISTER PUBLIC USER ID, PRIVATE USER ID, UE IP ADDRESS
85. Cx-PUT/ CX-PULL
86. Cx-PUT RESP/ Cx-PULL RESP
87. PUBLIC USER ID, PRIVATE USER ID, UE IP ADDRESS
CRF (OR PCCN ETC)
88. ACK
89. 2XX OK
90. 2XX OK
91. 2XX OK

Fig. 5

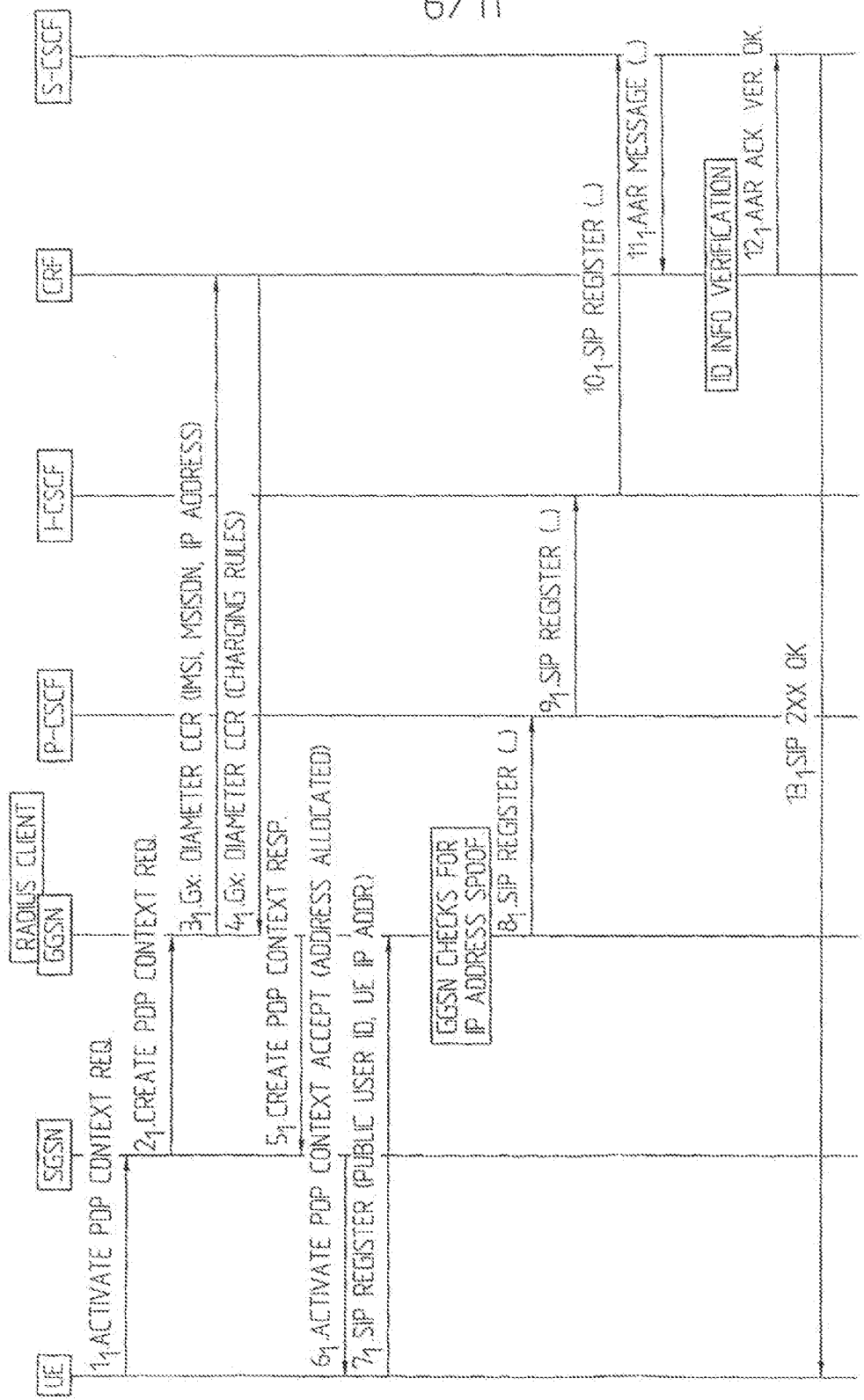

UE
SGSN
RADIUS CLIENT
GGSN
P-CSCF
I-CSCF
CRF
S-CSCF
1₁ ACTIVATE PDP CONTEXT REQ.
2₁ CREATE PDP CONTEXT REQ.
3₁ Gx: DIAMETER CCR (IMSI, MSISDN, IP ADDRESS)
4₁ Gx: DIAMETER CCR (CHARGING RULES)
5₁ CREATE PDP CONTEXT RESP.
6₁ ACTIVATE PDP CONTEXT ACCEPT (ADDRESS ALLOCATED)
7₁ SIP REGISTER (PUBLIC USER ID, UE IP ADDR)
GGSN CHECKS FOR IP ADDRESS SPOOF
8₁ SIP REGISTER (..)
9₁ SIP REGISTER (..)
10₁ SIP REGISTER (..)
11₁ AAR MESSAGE (..)
ID INFO VERIFICATION
12₁ AAR ACK VER. OK
13₁ SIP 2XX OK

Fig. 6A

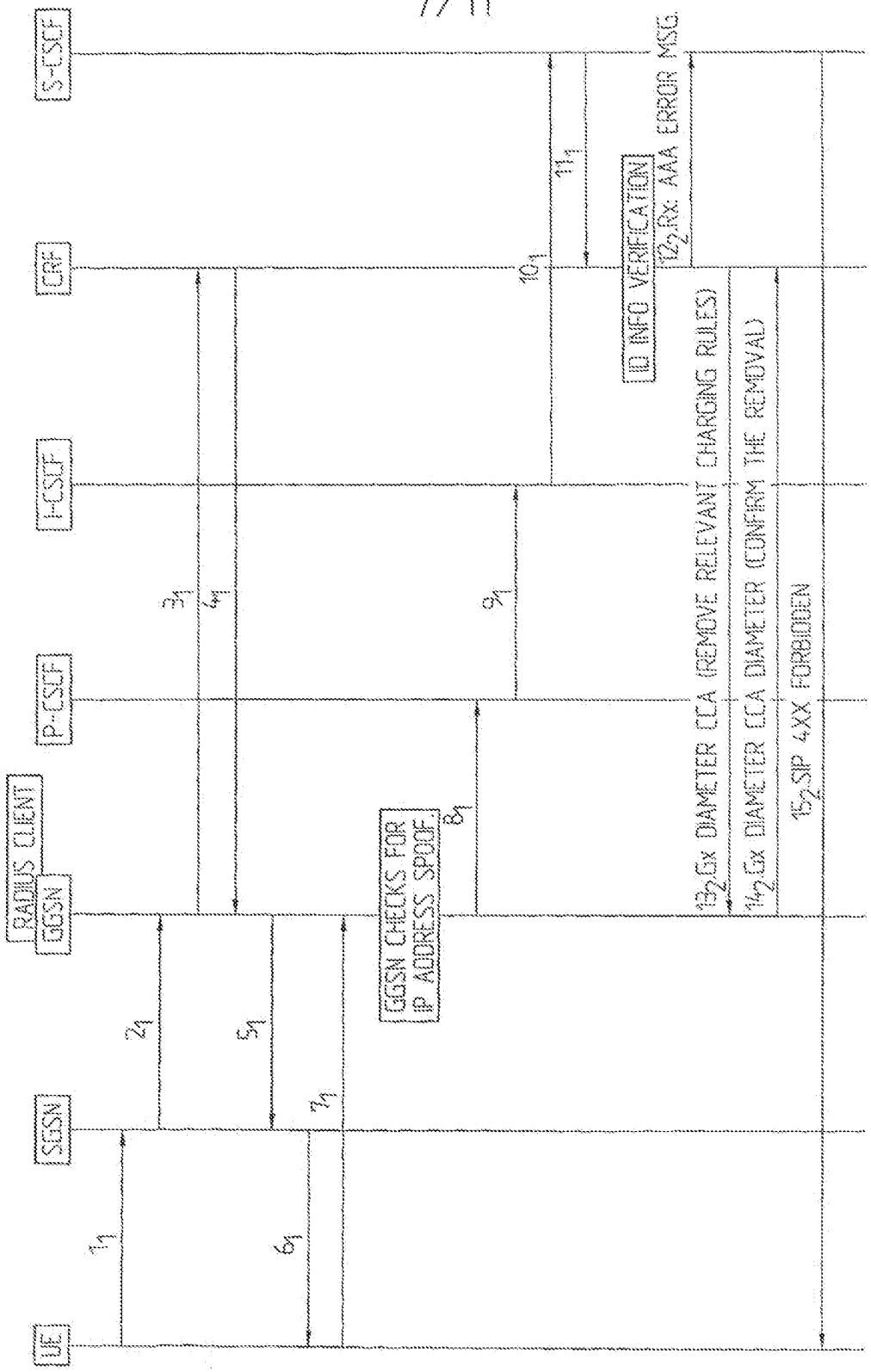

UE
SGSN
RADIUS CLIENT
GGSN
P-CSCF
I-CSCF
CRF
S-CSCF
1₁
2₁
3₁
4₁
5₁
6₁
7₁
8₁
9₁
10₁
11₁
12₂ Rx: AAA ERROR MSG
ID INFO VERIFICATION
13₂ Gx: DIAMETER CCA (REMOVE RELEVANT CHARGING RULES)
14₂ Gx: DIAMETER CCA (CONFIRM THE REMOVAL)
15₂ SIP 4XX FORBIDDEN
GGSN CHECKS FOR IP ADDRESS SPOOF

Fig. 6B

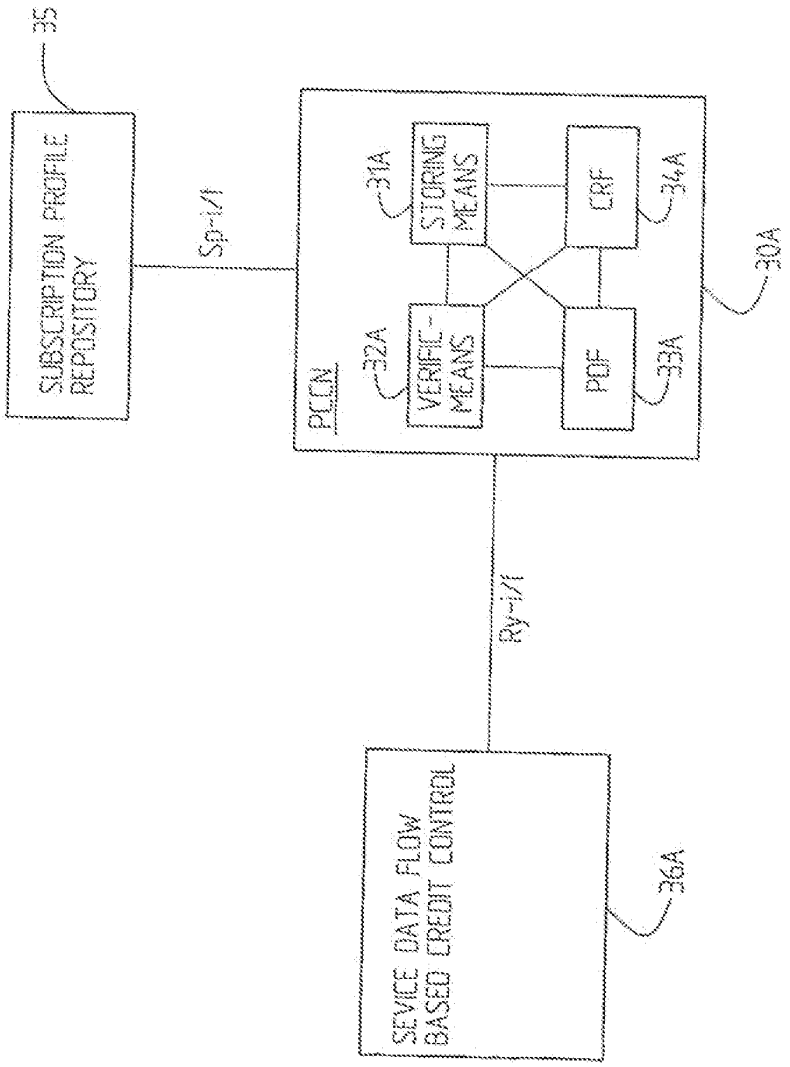

35
SUBSCRIPTION PROFILE REPOSITORY
Sp-i/f
PCEN
31A
STORING MEANS
CRF
34A
32A
VERIFIC-MEANS
PDF
33A
30A
Ry-i/f
SEVICE DATA FLOW BASED CREDIT CONTROL
36A

Fig. 7

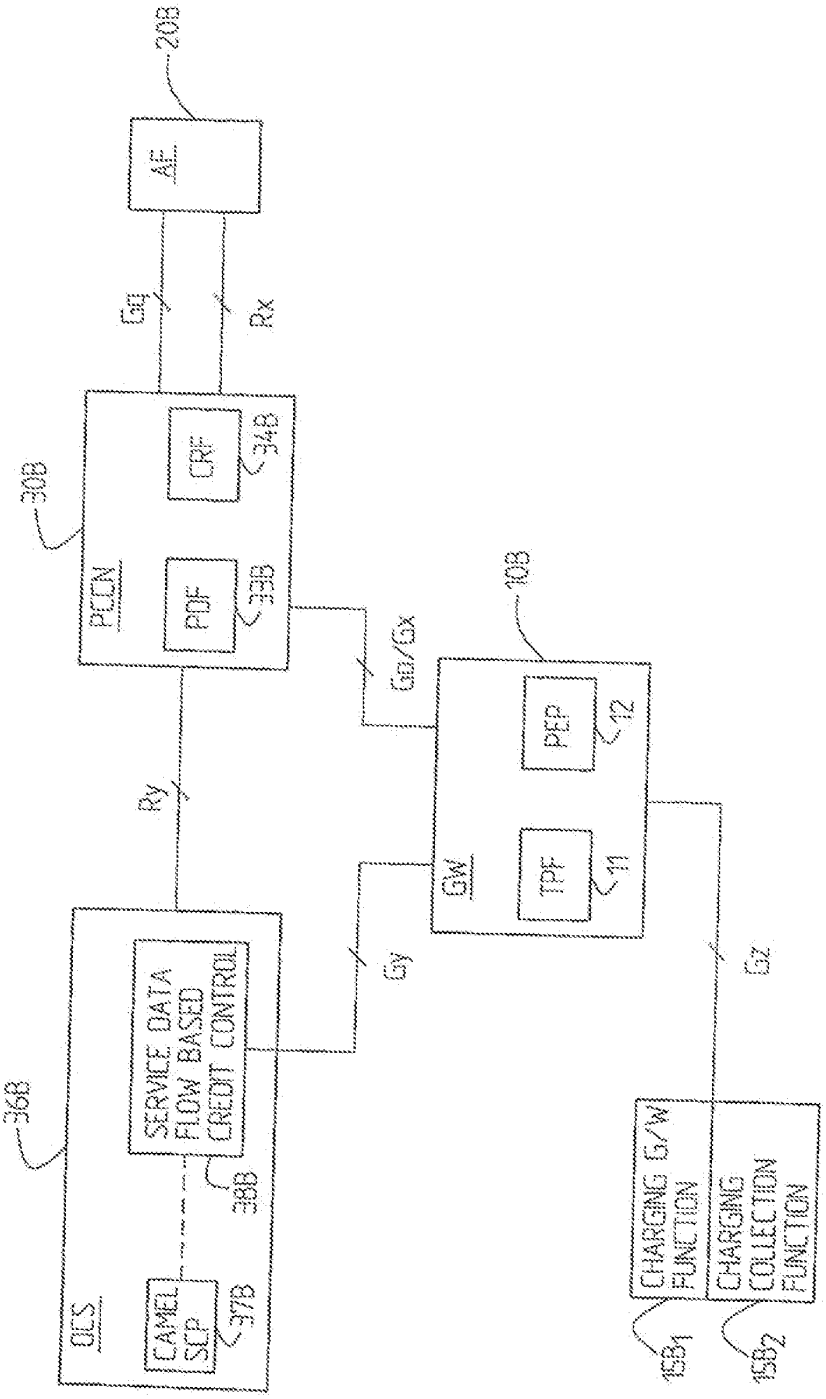

AF
20B
Rx
CRF
34B
PCEN
30B
PDF
33B
Ry
Gx/Gx
10B
PEP
12
GW
TPF
11
Gy
Gz
SERVICE DATA FLOW BASED CREDIT CONTROL
36B
38B
CAMEL SCP
37B
UES
CHARGING G/W FUNCTION
CHARGING COLLECTION FUNCTION
$1SB_1$
$1SB_2$

Fig. 8A

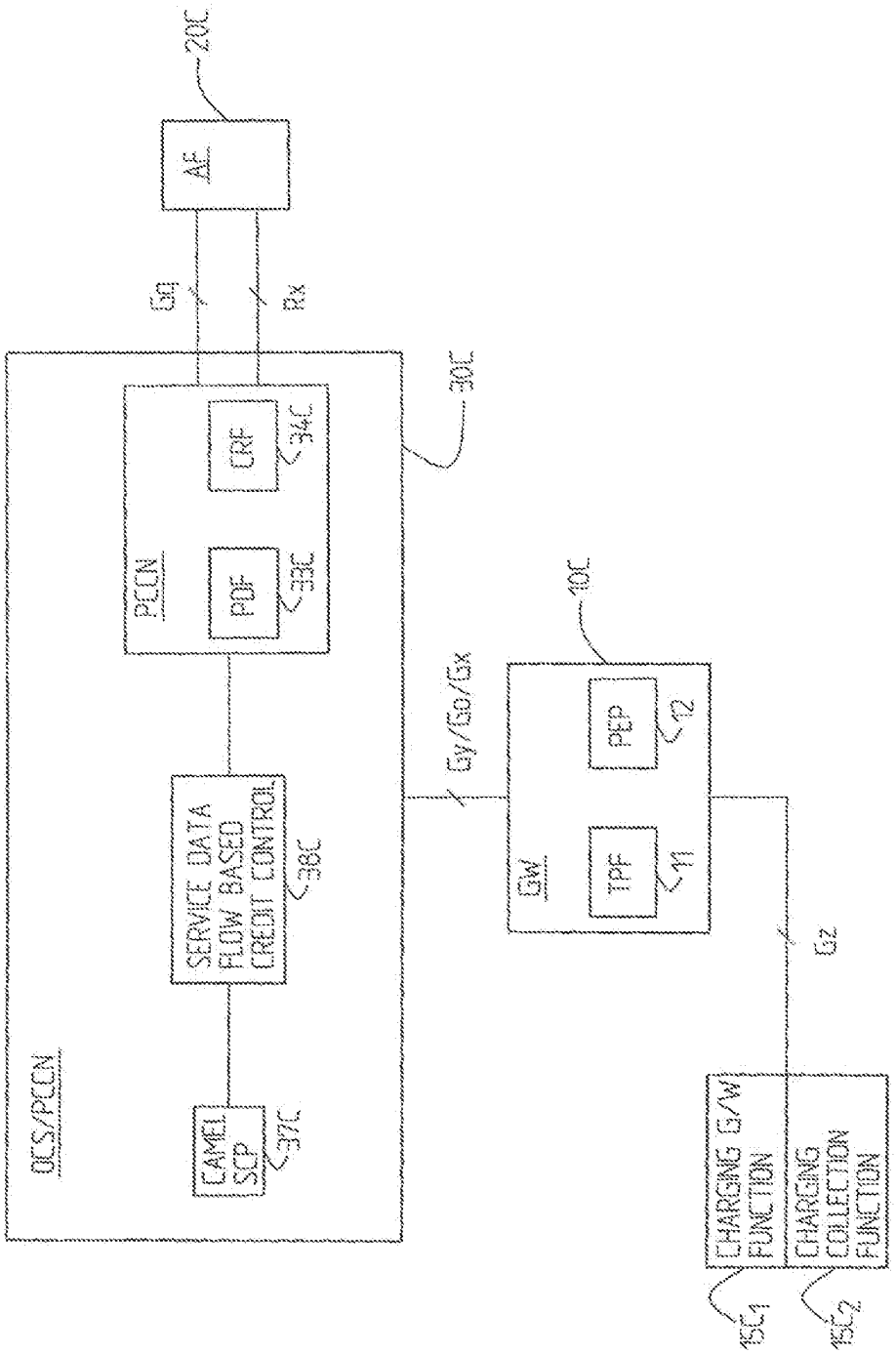

AF
20C
Gq
Rx
30C
PCRF
CRF
34C
PDF
33C
SERVICE DATA FLOW BASED CREDIT CONTROL
38C
CAMEL SCP
37C
OCS/PCRF
Gy/Gq/Gx
GW
10C
PEP
12
TPF
11
Gz
CHARGING G/W FUNCTION
CHARGING COLLECTION FUNCTION
15C1
15C2

Fig. 8B

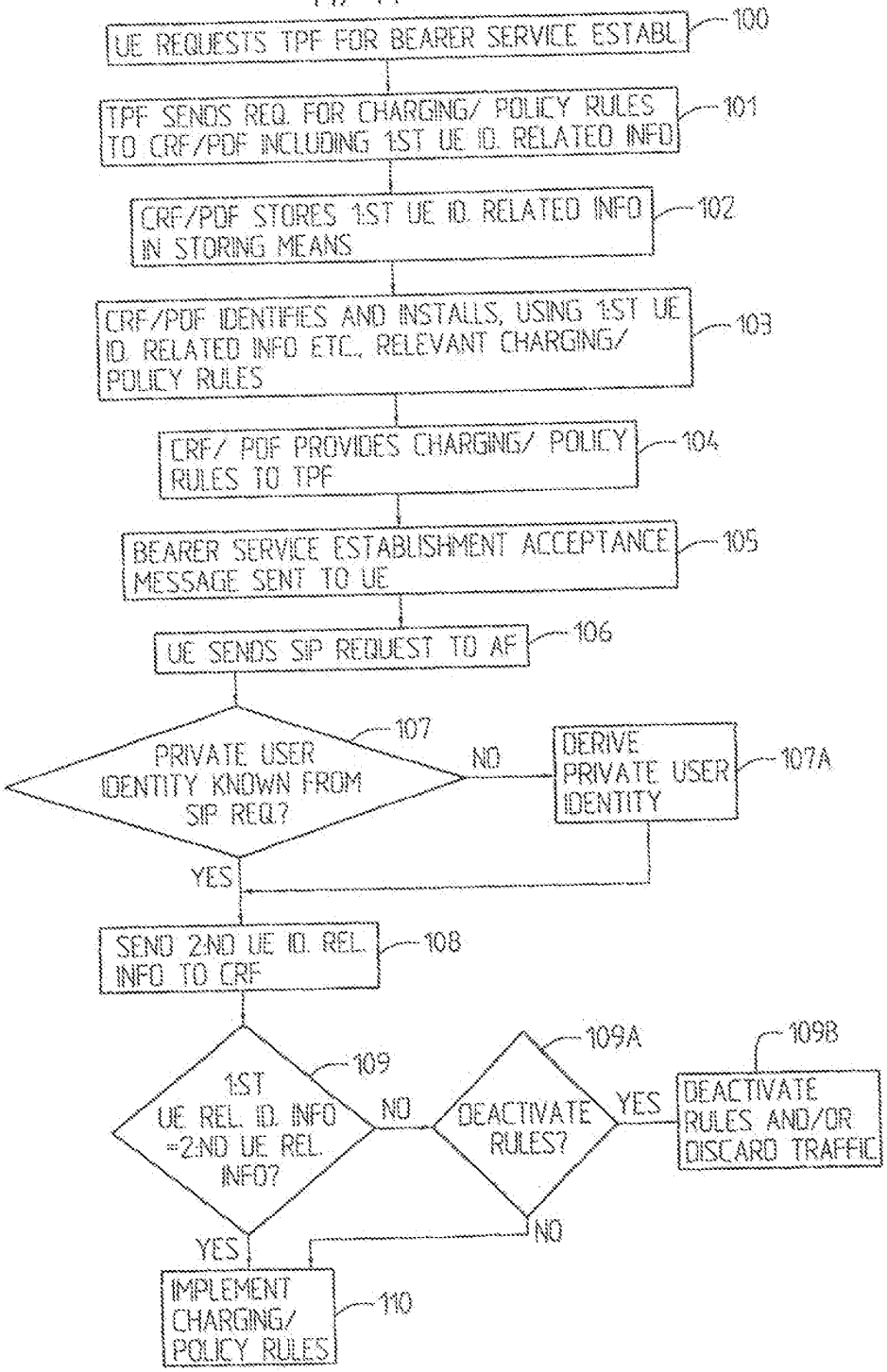

UE REQUESTS TPF FOR BEARER SERVICE ESTABL
100
TPF SENDS REQ. FOR CHARGING/ POLICY RULES TO CRF/PDF INCLUDING 1ST UE ID. RELATED INFO
101
CRF/PDF STORES 1ST UE ID. RELATED INFO IN STORING MEANS
102
CRF/PDF IDENTIFIES AND INSTALLS, USING 1ST UE ID. RELATED INFO ETC., RELEVANT CHARGING/ POLICY RULES
103
CRF/ PDF PROVIDES CHARGING/ POLICY RULES TO TPF
104
BEARER SERVICE ESTABLISHMENT ACCEPTANCE MESSAGE SENT TO UE
105
UE SENDS SIP REQUEST TO AF
106
PRIVATE USER IDENTITY KNOWN FROM SIP REQ?
107
NO
DERIVE PRIVATE USER IDENTITY
107A
YES
SEND 2ND UE ID. REL. INFO TO CRF
108
1ST UE REL. ID. INFO =2ND UE REL. INFO?
109
NO
DEACTIVATE RULES?
109A
YES
DEACTIVATE RULES AND/OR DISCARD TRAFFIC
109B
NO
YES
IMPLEMENT CHARGING/ POLICY RULES
110

Fig. 9

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2005/002837

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L29/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L H04M H04Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 01/91446 A (NOKIA CORPORATION; NOKIA INC) 29 November 2001 (2001-11-29) page 4, line 7 - page 5, line 2	1-38

[ ] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 June 2005	06/07/2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Veen, G

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP2005/002837

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0191446 A	29-11-2001	US 2002127995 A1	12-09-2002
		CN 1444824 A	24-09-2003
		EP 1310085 A2	14-05-2003
		WO 0191446 A2	29-11-2001
		JP 2004507127 T	04-03-2004

Form PCT/ISA/210 (patent family annex) (January 2004)

Electronic Acknowledgement Receipt

EFS ID:	8297275
Application Number:	12694445
International Application Number:	
Confirmation Number:	6030
Title of Invention:	SECURITY TECHNIQUES FOR DEVICE ASSISTED SERVICES
First Named Inventor/Applicant Name:	Gregory G. Raleigh
Customer Number:	21912
Filer:	Michael J. Schallop/Veronica Pula
Filer Authorized By:	Michael J. Schallop
Attorney Docket Number:	RALEP025
Receipt Date:	27-AUG-2010
Filing Date:	27-JAN-2010
Time Stamp:	16:04:48
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	RALEP025_IDS_Before_1st_OA.pdf	56034 b36609e32bb1573319e137a68cec19f5bb392ab0	no	2

Warnings:	
Information:	

2	Information Disclosure Statement (IDS) Filed (SB/08)	RALEP025_IDS_SB08_1.pdf	124560 04503647c440420ab8d5330a24057b2b76e1deea	no	4
Warnings:					
Information:					
This is not an USPTO supplied IDS fillable form					
3	Foreign Reference	RALE_Foreign_Ref_1A.pdf	6139391 ed22a6dbc67f09d4e7a5cb6045be867df9248ed8	no	49
Warnings:					
Information:					
4	Foreign Reference	RALE_Foreign_Ref_1B.pdf	2574136 928659e7726839520819 7f68febcdb55b01b915b	no	15
Warnings:					
Information:					
5	Foreign Reference	RALE_Foreign_Ref_1C.pdf	505198 81b92d38869e61daf2868357bed0ea81e29663a3	no	11
Warnings:					
Information:					
6	Foreign Reference	RALE_Foreign_Ref_1D.pdf	3048845 57a0856b7ec9abac7180d86595149dec996defdc	no	26
Warnings:					
Information:					
7	Foreign Reference	RALE_Foreign_Ref_1E.pdf	4616315 3e83c534eb55e82e478505c4d44bde2a72430854	no	34
Warnings:					
Information:					
8	Foreign Reference	RALE_Foreign_Ref_1F.pdf	5048671 cbad7159446ac414f1eeeed8214ef71e2adae8b4	no	38
Warnings:					
Information:					
9	Foreign Reference	RALE_Foreign_Ref_1G.pdf	5170774 c71740b2cd110174f5014e05d84acde040a43046	no	42
Warnings:					
Information:					
10	Foreign Reference	RALE_Foreign_Ref_1H.pdf	4124139 10eff1864a09668c5404e62fec900f25955b630d	no	36

Warnings:					
Information:					
11	Foreign Reference	RALE_Foreign_Ref_1I.pdf	3025460 60217d3f2867e38ce94f8f73ab1808b6938ba537	no	19
Warnings:					
Information:					
12	Foreign Reference	RALE_Foreign_Ref_1J.pdf	2632528 9240d5a44ab7e0daac9512d49a555fab8a37fb7d	no	13
Warnings:					
Information:					
13	Foreign Reference	RALE_Foreign_Ref_1K.pdf	7862034 663c74d022b4640fef56b6d06154d0afc06193e8	no	51
Warnings:					
Information:					
14	Foreign Reference	RALE_Foreign_Ref_1L.pdf	1929799 84c3744ddeee061cbb6a630d437dd3fc9d5f0a42	no	15
Warnings:					
Information:					
15	Foreign Reference	RALE_Foreign_Ref_1M.pdf	7699736 96ac86f4bbbc63af508f95f33927856e8c2f44c9	no	54
Warnings:					
Information:					
16	Foreign Reference	RALE_Foreign_Ref_1N.pdf	5831365 52c7919ee8f61002b63607e0b1afbb1c30411471	no	41
Warnings:					
Information:					
17	Foreign Reference	RALE_Foreign_Ref_1O.pdf	941474 84b1feb984dd76f2b40de9e9bf24bca6cbc317aa	no	18
Warnings:					
Information:					
18	Foreign Reference	RALE_Foreign_Ref_1P.pdf	4880569 cf0b517f0ea17a760054e4b34ec60d67a2caf375	no	37
Warnings:					
Information:					
19	Foreign Reference	RALE_Foreign_Ref_1Q.pdf	5439339 547fe6d96444a8d3f0953a82899a468dcaa3c29b	no	37

Warnings:					
Information:					
20	Foreign Reference	RALE_Foreign_Ref_1R.pdf	7122128 4ad5c3416c210e3f1efbd1c3e886d2ae2ad8e879	no	52
Warnings:					
Information:					
			Total Files Size (in bytes):	78772495	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor:	Gregory G. Raleigh	Examiner:	Not Assigned
Application No.:	12/694,445	Art Unit:	2614
Filed:	January 27, 2010	Docket No.:	RALEP025
Title:	SECURITY TECHNIQUES FOR DEVICE ASSISTED SERVICES		

INFORMATION DISCLOSURE STATEMENT
UNDER 37 CFR §§1.56 AND 1.97(b) BEFORE MAILING OF A FIRST OFFICE ACTION

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

The references listed in the attached PTO SB/08, may be material to examination of the above-identified patent application. Applicants disclose these references in compliance with their duty of disclosure pursuant to 37 CFR §§1.56 and 1.97. The Examiner is requested to make these references of official record in this application. Copies of the cited references are provided to the extent required by 37 CFR §1.98 and any applicable waivers.

This Information Disclosure Statement is not to be construed as a representation that a search has been made, that additional information material to the examination of this application does not exist, or that these references indeed constitute prior art.

This Information Disclosure Statement is being filed (1) with the above-referenced application or (2) within three months of the filing date of the above-referenced application or (3) before the mailing date of a first Office Action on the merits or (4) before the mailing date of a first Office Action after filing of a Request for Continued Examination. Accordingly, it is believed that no fees are due in connection with the filing of this Information Disclosure Statement. However, if it is determined that any fees are due, the Commissioner is hereby authorized to charge such fees to Deposit Account 50-0685 (Order No. RALEP025).

Respectfully submitted,
VAN PELT, YI & JAMES LLP

/Michael J. Schallop/

Michael J. Schallop
Registration No. 44,319
V (408) 207-4762
F (408) 973-2595

10050 N. Foothill Blvd, Ste 200
Cupertino, CA 95014

Substitute for form 1449/PTO				Complete if Known	
Information Disclosure Statement By Applicant *(Use as many sheets as necessary)*				Application No.:	12/694,445
				Filing Date	January 27, 2010
				First Named Inventor	Gregory G. Raleigh
				Art Unit	2614
				Examiner Name	Not Assigned
Sheet	1	of	1	Atty Docket No.	RALEP025

U.S. Patent Documents					
Examiner Initials	CiteNo.	Document Number Number-Kind Code (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	3AA	US 7,747,240	06-29-2010	BRISCOE et al.	
	3BB	US 2003/0088671	05-08-2003	KINKER et al.	
	3CC	US 2007/0248100	10-25-2007	ZUBERI et al.	
	3DD	US 2008/0130534	06-05-2008	Tazuko Tomioka	

FOREIGN PATENT DOCUMENTS					
Examiner Initials	Cite No.	Foreign Patent Document Country Code–Number–Kind Code	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	3A	WO 2007/014630	02-08-2007	Telefonaktiebolaget LM Ericsson	
	3B	WO 2006/004467	01-12-2006	Telefonaktiebolaget LM Ericsson	
	3M	WO 1999/065185	12-16-1999	British Telecommunications	

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published	T^2
	3C	KNIGHT et al., Layer 2 and 3 Virtual Private Networks: Taxonomy, Technology, and Standardization Efforts, IEEE Communications Magazine, June 2004	
	3D	KOUTSOPOULOU et al., Middleware Platform for the Support of Charging Reconfiguration Actions, 2005	
	3E	NILSSON et al., A Novel MAC Scheme for Solving the QoS Parameter Adjustment Problem in IEEE802.11e EDCA, February 2006	
	3F	Author Unknown, Overview of GSM, GPRS, and UMTS, Chapter 2, November 4, 2008	
	3G	CHAOUCHI et al., Policy Based Networking in the Integration Effort Of 4G Networks and Services, 2004 IEEE	
	3H	ZHU et al., A Survey of Quality of Service in IEEE 802.11 Networks, IEEE Wireless Communications, August 2004	
	3I	KYRIAKAKOS et al., Ubiquitous Service Provision in Next Generation Mobile Networks, Proceedings of the 13th IST Mobile and Wireless Communications Summit, Lyon, France, June 2004	
	3J	FAROOQ et al., An IEEE 802.16 WiMax Module for the NS-3 Simulator, March 2-6, 2009	
	3K	Author Unknown, HP, IP Multimedia Services Charging, A White Paper from HP, January 2006	
	3L	Author Unknown, Kindle™ User's Guide 3rd Edition, Copyright 2004-2009	

Examiner	Date Considered

Examiner: Initial citation considered. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T^2 = Applicant is to place a check mark here if English language Translation is attached

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
8 February 2007 (08.02.2007)

(10) International Publication Number
WO 2007/014630 A1

(51) International Patent Classification:
H04Q 7/38 (2006.01) *H04M 3/42* (2006.01)
H04L 29/08 (2006.01)

(21) International Application Number:
PCT/EP2006/006776

(22) International Filing Date: 11 July 2006 (11.07.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
05016857.4 3 August 2005 (03.08.2005) EP

(71) Applicant *(for all designated States except US)*: **TELEFONAKTIEBOLAGET LM ERICSSON (publ)** [SE/SE]; S-164 83 Stockholm (SE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **NORRMAN, Karl** [SE/SE]; Bondegatan 3b, 4 Tr, S-116 23 Stockholm (SE). **BARRIGA, Luis** [SE/SE]; Glidflyktsgatan 4 Bv, S-128 33 Skarpnäck (SE). **HALEN, Joacim** [SE/SE]; Anhaltsvägen 8, S-191 40 Sollentuna (SE).

(74) Agent: **HOFFMANN . EITLE**; Arabellastrasse 4, 81925 München (DE).

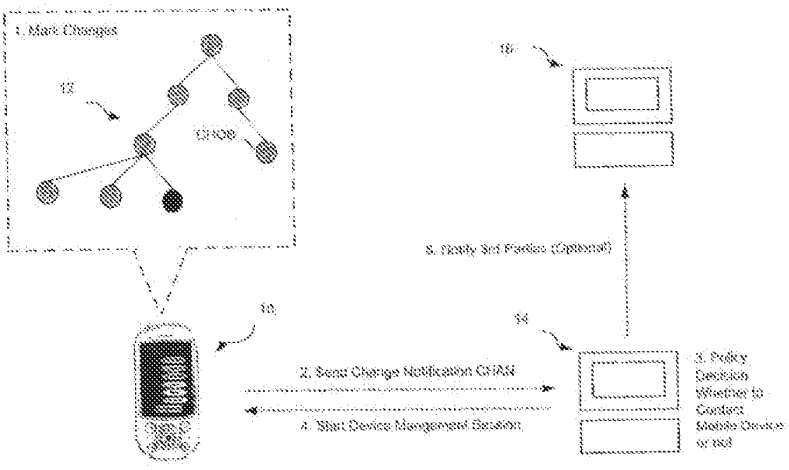
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: AUTOMATIC MOBILE DEVICE CAPABILITY MANAGEMENT



(57) Abstract: The present invention relates to an improved approach to mobile device capability management. Heretofore, a capability management device (14) is provided at a mobile communication network. Upon change of a mobile device capability, a related notification is sent to the capability management device (14) which applies a policy decision whether to track the capability change at the network side or not. Should the decision be 'yes', the capability management device (14) starts a device management session to collect further information on a mobile device capability change beyond the information made available with the mobile device capability change notification. After retrieval of the mobile device capability change information, the capability management device (14) will update its mobile device capability state accordingly. Optionally, the capability management device (14) may share the generated information with third party (16) cooperating with the mobile device (10) for, e.g., service delivery.

WO 2007/014630 A1

Automatic Mobile Device Capability Management

FIELD OF INVENTION

The present invention relates to a method of operating a mobile device capability management apparatus in a mobile communication network and also to a method of operating a mobile device being enabled for mobile device capability management. Further, the present invention relates to a corresponding mobile device capability management apparatus and a mobile device interacting with the mobile device capability management apparatus.

BACKGROUND ART

As mobile devices are becoming more complex, flexible and capable of more functionality, each mobile device evolves into an increasingly unique user specific equipment diverging from previously pre-installed functionalities. In the telecommunication world, this leads to a big change from earlier days, when mobile devices of a certain model were static with respect to both hardware and to a large extent even software.

Today, mobile devices communicate their capabilities to a wireless communication network when attaching thereto by sending its identification ID and the version number reflecting its entire software image. Further, the operator

of the wireless communication network has the option, upon reception of the mobile device identification ID, to retrieve a stored version of the software installed in the mobile device into a mobile communication network operator controlled database.

Fig. 1 shows a sample management tree maintained in support of mobile device capability management.

One specification of such a management tree is provided by www.openmobilealliance.com and specified in OMA, SyncML device management tree and description, version 1.1, February 2002.

As shown in Fig. 1, the management tree serves to specify the structure of the configuration capability present in a mobile device. The main idea is to use the management tree, which resembles, e.g., a UNIX file system. The management tree contains sub-trees called management objects MO. These management objects are composed of information required to form a specific task, e.g., they may be related to device specific information, MMS functionality or email functionality. A device information management object contains a leaf node named './DevInfo/Lang', holding information on the currently selected operative language for the mobile device.

As shown in Fig. 1, leaf nodes may be identified using a relative unified resource identifier URI, so that it is possible to read and write content of leaf nodes using standards like Synchronization Meta language SyncML. Also, such synchronization meta languages may be used to add and remove nodes in the management tree.

Further, while it is possible for the operator to update the information in the leaves of the management tree while the mobile device is attached to the mobile communication network, there exists also the possibility to change settings and to upgrade the functionality of the mobile device from other resources than the operator of the mobile communication network. Also, this may be achieved when the mobile device is not attached to the mobile communication network. Should such a situation occur, the operator would be interested in knowing what has changed in the mobile device in order to provide such information to, e.g., service applications interacting with the mobile device or third party service providers that deliver device and their content and services to the mobile device.

Here, when a new service is made available in a mobile device, this will typically lead to a new management object being installed in the management tree. In many cases, e.g., for reasons of consistency, it would be interesting for the operator of the mobile communication network to retrieve also these extensions to the management tree. It should be noted that it is not always possible for the operator to know such information in advance.

An additional problem with existing solutions is that information conveyed by the mobile device on its capabilities is usually too coarse. As outlined above, existing solutions only provide a software version number for the entire software image present in the mobile device. Typically, such information could be sent from the access network, MSC/HLR, to a service layer device management device using automatic device detection procedures, as an example.

SUMMARY OF INVENTION

In view of the above, the object of the present invention is to provide an improved management of mobile device capabilities.

According to the present invention, this object is achieved by a method of operating a mobile device capability management device in a mobile communication network with respect to at least one mobile device served by the mobile communication network. According to the present invention, a mobile device capability is represented two-fold, i.e. as a first mobile device profile in the mobile device capability management device operated by a network operator and as a second mobile device profile maintained in the mobile device. Upon change of mobile device capability and functionality, at the mobile device capability management device, there is received a mobile device capability change information with respect to the mobile device capability change.

According to the present invention, it is suggested that the mobile device capability management device executes a step to assess the need to update the first mobile device profile in the mobile device capability management device. This assessment is achieved on the basis of a mobile device capability management policy and the received mobile device capability change information.

Therefore, the present invention provides the possibility for a mobile device capability management system run by an operator of a mobile communication network to get up-to-date information on the mobile device capability change and therefore a detailed view on the mobile device capabilities in cases where the mobile device was updated via some

facilities being different from the mobile device capability management device. Contrary to existent solutions, where only the version number of the software image is conveyed to an operator of a mobile wireless network, the present invention allows for a much finer granularity of information sent from the mobile device to a mobile device capability management device.

Further, the invention enables efficient timely automatic detection of mobile device capability changes like firmware changes or device software changes, and to retrieve information on such changes, even with respect to specific components and functionalities, e.g., an updated audio or video codec, newly installed/updated applications, and/or change of user preferences, e.g., preferred language.

According to a preferred embodiment of the present invention, there is provided the step of actually updating the mobile device profile in the mobile device capability when the assessment step indicates necessity of such an update.

Yet another important advantage of the present invention is that it is possible to execute the assessment step in a very flexible way. In other words, according to a specified mobile device capability management policy, one could, e.g., decide to track every mobile device capability change in the mobile device capability management device, or to only track such changes according to specific time windows, or with respect to certain functionalities, or according to subscriber information, whatsoever. The option to specify a mobile device capability management policy is highly advantageous to avoid unnecessary steps at the mobile device capability management device, should only minor changes occur with respect to the capability of the mobile

device. The avoidance of such unnecessary update steps at the mobile device capability management device is also a prerequisite to avoid unnecessary traffic in the mobile communication network run by the operator, which also runs the mobile device capability management device.

In more detail, according to a preferred embodiment, only upon assessment of a necessary update, there is established a mobile device management session between the mobile device capability management device and the mobile device for evaluation of the second mobile device profile in the mobile device. Then, related mobile device capability update information for processing in the mobile device capability management device is received. Preferably, such information may be provided as a mobile device management object being specified for a mobile device management tree structure which may be dedicated to the mobile device capability change management.

An important advantage of this procedure is that, as outlined above, overall mobile device management traffic can be minimized according to the established mobile device capability management policy, and that further mobile device capability change information is transferred in a way suitable for a mobile device management tree data structure used in the mobile device capability management device.

According to a preferred embodiment of the present invention, the mobile device capability change information contains a qualitative parameter indicating seriousness of change to the mobile device profile and a quantitative parameter indicating amount of detective changes to the mobile device profile.

Typically, seriousness of change reflects a degree of change or, in other words, minor, medium, major, as an example. This information is helpful to assess at the mobile device capability management device whether to actually establish a mobile device management session or not. Further, the amount of detective changes reflects the actual changes added to the mobile device profile and therefore gives information on what type of functionality has actually been changed at the mobile device side.

This two-step approach to forwarding of mobile device capability change information is very effective in that in a first step only a degree of change is indicated to the mobile device management device and then, in a second step, detailed information necessary for mobile device capability profile update at the mobile device capability management device is exchanged, e.g., by establishment of a mobile device capability management session.

According to a preferred embodiment of the present invention, the mobile device capability management device further notifies a service provider delivering services to the mobile device and/or an application interacting with the mobile device on a mobile device capability change.

An important advantage of this notification is that not only at the operator side running the mobile communication environment information reflecting a mobile device profile update is available, but also at third parties. This spread out of mobile device capability change information to all parties interacting with the mobile device achieves consistency throughout the complete operative environment and therefore an increase in quality of service delivery.

Further to the above, the object outlined above is also achieved according to a second aspect of the present invention by a method of operating a mobile device being enabled for mobile device capability management through a mobile communication network. Here, in the mobile device a mobile device capability state is represented as a mobile device profile.

According to the present invention, the method of operating a mobile device comprises the step of monitoring a change in the mobile device capability. Then, should a change in the mobile capability be monitored, there is generated a mobile device capability change information with respect to a mobile device capability change. Then, the mobile device executes a step to forward the generated mobile device capability change information to the mobile communication network. Therefore, an important advantage of the mobile device according to the present invention is that it is not only adapted to receive instructions for mobile device capability management, but that it also provides functionality in itself to continuously monitor the mobile device capability. This monitoring is a prerequisite for initiation of a capability change notification to the mobile communication network for improved mobile device capability management therein. Therefore, the operator may effectively and in real-time execute a mobile device capability update in support of back-up services for end users and/or in support of mobile device capability sharing towards external mobile device relevant applications.

According to a preferred embodiment, the mobile device capability change information comprises a qualitative parameter indicating seriousness of change to the mobile device profile and a quantitative parameter indicating amount of effective changes to the mobile device profile.

Similar to the advantages given above, the mobile device is adapted to support a two-stage approach to capability change information delivery so as to reduce traffic for mobile device capability management.

Also, it should be mentioned that according to the present invention, mobile device capability change information may be exchanged with the mobile communication network either through a circuit switched mobile communication, a packet switched mobile communication and/or through service layer data exchange or any hybrid form thereof.

Again, the applicability of different types of mobile communication support flexibility and availability of information exchange during mobile device capability management.

According to another preferred embodiment of the present invention there is provided a computer program product directly loadable into the internal memory of a mobile device capability management device in a mobile communication network comprising software code portions for performing the inventive mobile device capability management process when the product is run on a processor of the mobile device capability management device.

According to another preferred embodiment of the present invention there is provided a computer program product directly loadable into the internal memory of a mobile device being enabled for mobile device capability management through a mobile communication network comprising software code portions for performing the

inventive mobile device capability update process when the product is run on a processor of the mobile device.

Therefore, the present invention is also provided to achieve an implementation of the inventive method steps on computer or processor systems. In conclusion, such implementation leads to the provision of computer program products for use with a computer system or more specifically a processor comprised in, e.g., a mobile device capability management device or a mobile device.

This programs defining the functions of the present invention can be delivered to a computer/processor in many forms, including, but not limited to information permanently stored on non-writable storage media, e.g., read only memory devices such as ROM or CD ROM discs readable by processors or computer I/O attachments; information stored on writable storage media, i.e. floppy discs and harddrives; or information convey to a computer/processor through communication media such as network and/or Internet and/or telephone networks via modems or other interface devices. It should be understood that such media, when carrying processor readable instructions implementing the inventive concept represent alternate embodiments of the present invention.

BRIEF DESCRIPTION OF DRAWING

In the following the best mode of carrying out the present invention and corresponding preferred embodiments will be described with reference to the drawing in which:

Fig. 1 shows a sample management tree maintained in support of mobile device capability management;

Fig. 2 shows a basic concept underlying the present invention;

Fig. 3 shows a schematic diagram of a mobile device according to the present invention and a related flowchart of operation;

Fig. 4 shows a schematic diagram of a mobile device capability management device according to the present invention and a related flowchart of operation;

Fig. 5 shows a further detailed flowchart of operation for the mobile device shown in Fig. 3;

Fig. 6 shows a further detailed schematic diagram of the mobile device capability management device shown in Fig. 4;

Fig. 7 shows a detailed flowchart of operation for the mobile device capability management device shown in Fig. 6;

Fig. 8 shows an example of marking changed nodes in a mobile device management tree;

Fig. 9 shows a further example of marking changed nodes in a mobile device capability management tree; and

Fig. 10 shows examples of paths for conveying mobile device capability change information between the mobile device and the mobile device capability management device.

DESCRIPTION OF BEST MODE AND PREFERRED EMBODIMENTS

In the following, the present invention will be described with respect to what is considered as best mode and preferred embodiments thereof. While in the following different principles underlying the present invention will be explained with respect to a mobile device capability management tree, nevertheless, it should be noted that the application of such a mobile device capability management tree is not to be considered as restricting the scope of the present invention, as long as any type of information reflecting the capability of the mobile device is maintained. Thus, it should be understood that any other

type of data structure suitable for maintaining a mobile device capability profile is well covered by the present invention.

Further, insofar as different functionality of the present invention is explained, it should be clear that any such functionality may be implemented in hardware or software and/or a combination thereof.

Fig. 2 shows a basic concept underlying the present invention.

As shown in Fig. 2, in a mobile device 10, there is maintained some type of information reflecting the mobile device capability which is shown in the form of a mobile device capability management tree 12 in Fig. 2. This structure forms the pre-requisite for implementing the present invention at the mobile device side, i.e. the execution of a marking phase. Here, during changes to the mobile device capability, the mobile device monitors such changes, collects information on changes and sets up mobile capability change information that may be conveyed to the mobile device capability management device, e.g., a device management server.

As shown in Fig. 2, a second element of the inventive mobile device management is a so-called notification phase, wherein the mobile device notifies the mobile device capability management device 14 on the device capability change(s). Preferably, this notification may include the

type of change and, optionally, a descriptor of the involved mobile device capability changes.

As shown in Fig. 2, a third element of the inventive mobile device management process is a so-called reaction phase, where the mobile device capability management device 14 receives the notification of a mobile device capability change and reacts accordingly. Such reaction is based on the received mobile device capability change information and a mobile device capability management policy. Here, it should be noted that such a policy reflects how accurate the changes of mobile device capability must be tracked at the mobile device capability management device, on what mobile device capability management traffic on the network would be acceptable, type of change and related handling of the mobile device capability management device 14, etc. Also, it should be noted that according to the present invention, such a mobile device capability management policy may either be static and predetermined or dynamically changed over time.

As shown in Fig. 2, a positive assessment on necessity of a mobile device capability profile update at the mobile device capability management device 14, the mobile device capability management device 14 will react through establishing a device management session with the mobile device 10 for subsequent receipt of information being related to the mobile device capability change in any appropriate form. Optionally, the mobile device capability management device 14 may notify third parties of the mobile

device capability change, once the related information is received from the mobile device 10.

Therefore, the basic concept underlying the present invention allows for a finer granularity of information being related to a mobile device capability change which is sent from the mobile device 10 to the mobile device capability management device 14. Contrary to existing solutions, only a version number of a software image is conveyed to the operator of the mobile communication network. The present invention enables automatic detection of mobile device capability changes and subsequent retrieval of information on changes made to specific components, e.g., an updated codec, an installed application, and/or change of the user preference such as a preferred language. Therefore, the operator of the mobile communication network can use this detailed knowledge on the mobile device capability for device services for end users and/or for mobile device capability sharing towards third parties 16 and further applications.

In the following, further details of the present invention will be explained with respect to Figs. 3 to 10. Here, Fig. 3 shows a schematic diagram of a mobile device according to the present invention and a related flowchart of operation.

As shown in Fig. 3, the mobile device 10 comprises a mobile device memory unit 18, a mobile device profile monitoring unit 20, a profile changing information generating unit 22, and a communication unit 24.

Operatively, the mobile device profile monitoring unit 20 executes a step S10 to monitor a change in the mobile device capability. Further, operatively the mobile change information generating unit 22 executes a step S12 to generate mobile device capability change information with respect to a mobile device capability change. Then, the communication unit 24 executes a step S14 to forward the mobile device capability change information to the mobile communication network. Any type of information reflecting the mobile device capability is stored in the mobile device profile memory unit 18.

Fig. 4 shows a schematic diagram of a mobile device capability management device 14 according to the present invention and a related flowchart of operation.

As shown in Fig. 4, the mobile device capability management device 14 comprises a communication unit 26, an assessment unit 28, and a mobile device profile updating unit 30.

As shown in Fig. 4, operatively, the communication unit 26 executes a step S16 to receive a mobile device capability change information with respect to a mobile device capability change from a mobile device 10 being enabled for mobile device capability management. Then, operatively the assessment unit 28 will execute a step S18 to assess the need to update the mobile device profile in the mobile device capability management device. This assessment is based on a mobile device capability management policy and the received mobile device capability change information. The mobile device capability management policy, either

static or changed dynamically over time, may reflect any criteria of value to the operator of the mobile communication environment, e.g., amount of desired mobile device management traffic, degree of changes at the mobile device side, e.g., minor, middle, major, focus on specific types of functionality updates at the mobile device side of particular relevance to the operator, e.g., update of speech coding software or lower layer communication software, etc. As shown in Fig. 1, should the assessment step S18 lead to a result that no update of a mobile device profile is necessary at the time being at the mobile device capability management device 14, then the operation of the mobile device capability management device 14 stops. Otherwise, the mobile device profile updating unit 30 will react to the information received with respect to a mobile device capability change.

Fig. 5 shows a further detailed flowchart of operation of the mobile device 10 being enabled for mobile device capability management shown in Fig. 3.

As shown in Fig. 5, the mobile device profile monitoring unit 20 shown in Fig. 3 continuously executes the step S10 to monitor changes in mobile device capability. It may be implemented as a set of functions that are initiated upon a change of a mobile device profile. After step S10 the mobile device profile monitoring unit 20 executes a step S20 to check whether a change in capability in detail has occurred. If this is not the case, it will branch back to the monitoring step S10 to repeat the monitoring of mobile device capability changes. Otherwise, the profile change

information generating unit 22 executes the step S12 to generate the mobile device capability change information as outlined above. Here, without restricting the scope of the present invention, one may assume that the mobile device profile is represented as a mobile device capability management tree. Further, according to the present invention, it is suggested that the management tree contains an object called 'changes object' CHOB. For each capability change in the mobile device 10, therefore the related changes object in the management tree will be modified to reflect the changes to the mobile device capability into the changes object. From the viewpoint of a management tree representation, the changes object may reflect, when another object is changed, added or deleted, or when a change to a leaf node of the management tree has happened, as will be explained in more detail with respect to Fig. 8 and 9, respectively.

As shown in Fig. 5, after generation of the mobile device capability change information, e.g., related update of the changes object CHOB, then the communication unit 24 will execute the step S14, as soon as the mobile device has network connectivity, to forward mobile device capability change information to the mobile device capability management device 14. This may be achieved by set-up of a modification notification CHAN message. Preferably, the change notification CHAN message consists of two parts:

- A *qualitative* parameter indicating the degree of changes to the management tree. E.g., the qualitative parameter may be qualified into major, medium or

minor. Otherwise, one may consider a qualitative parameter being scaled from 0 to 1 to have a finer grid of indicating of the amount of changes.

- The second parameter is a *quantitative* parameter indicating the amount of effective changes to the management tree. Here, one may consider, e.g., a situation where an end user modified, e.g., a speech codec of a mobile device 10. Assuming that the speech codec is indeed activated, the quantitative parameter would indicate one changed node in the management tree. Otherwise, assuming that the speech codecs envisaged by the end user is not compatible with the operative system of the mobile device 10 and would therefore not be activated, the quantitative parameter would indicate an amount of 0 with respect to this change, notwithstanding the further amendments to the management tree, which may be notified at the same time with the change notification CHAN message.

As shown in Fig. 5, further steps of operation of the mobile device 10 may relate to the interaction with the mobile device capability management device 14 shown in Fig. 4.

As shown in Fig. 5, the communication unit 24 of the mobile device 10 may execute a step 822 to receive a request for mobile device capability change information from the mobile communication network, which information would be more detailed over the change notification CHAN message referred to above. Then, the communication unit 24 executes a step

S24 to establish a mobile device capability management communication session with the mobile device capability management device 14.

Then, the communication unit 24 will execute a step S26 to forward detailed mobile device capability change information to the mobile device capability management device 14. This detailed information is set up by the profile change information generating unit in cooperation with a mobile device profile memory unit 18 and the mobile device profile monitoring unit 20.

Fig. 6 shows a further detailed schematic diagram of the mobile device capability management device 14 explained with respect to Fig. 4.

As shown in Fig. 6, the mobile device profile updating unit 30 shown in Fig. 4, according to the present invention, divides into a session initiation unit 32, a change information processing unit 34, and a mobile device profile memory 36.

Fig. 7 shows a further detailed flowchart of operation for the mobile device capability management 14 shown in Fig. 6.

As shown in Fig. 7, the communication unit 26 will execute a step S28 to monitor whether a mobile device capability change information, e.g., a change notification message CHAN, has been received at the mobile device capability management device 14. If so, the assessment unit 28 will execute the step S18 to assess the need to update a mobile

device profile in the mobile device capability management device 14. In more detail, when the mobile device capability management device 14 is in receipt of a change notification message CHAN, the assessment unit 28 will retrieve the device mobile stored in the mobile device profile memory 36. Assessing the quantitative and qualitative indicators in the change notification message CHAN in relation to the retrieved device profile, the assessment unit 28 may assess the need to contact the mobile device 10 for related mobile device profile update, e.g., by modifying the management tree with respect to the mobile device 10.

E.g., in the assessment step S18, there is only a minor change to the mobile device capability. The mobile device capability management device 14 may decide to not contact the mobile device 10. However, should there be changes to critical objects in the management tree of the mobile device which touch operator policies or third party service providers, then the operator may choose to scan the mobile device 10 as soon as possible.

As shown in Fig. 7, heretofore the communication unit 26 will execute a step S32 to establish a mobile device management session with a mobile device 10 for retrieval of further detailed information on a mobile device capability change. The same communication unit 26 will execute a step S34 to receive detailed mobile device capability change information from the mobile device 10, e.g., according to a changes object CHOB. Should the mobile device capability management device 14 decide to contact the mobile device

10, it should then receive information on the CHOB object and the management tree itself in order to obtain necessary details about changes done to the management tree. In view of the above, it becomes clear that the approach according to the present invention may be basically divided into a marking phase, a notification phase, and a reaction phase. In particular, qualifiers used within the change notification CHAN message are used to prevent the mobile device capability management device from scanning the whole mobile device 10 for changes, each time the mobile device 10 is upgraded without involvement of the mobile device capability management device 14. Therefore, the present invention avoids time-consuming, unnecessary update procedures, disturbance of normal mobile device operation, generation of heavy traffic in the mobile communication network, and in particular with respect to mobile communication where cellular bandwidth is a very limited resource.

Then, the change information processing unit 34 will execute a step S36 to update the mobile device profile, e.g., a management tree, according to the received information and initiate a related update of the corresponding information in the mobile device profile memory 36. Optionally, the change information processing unit 34 may execute a step S38 to notify an information on a mobile device profile update to a service provider or in application delivery services.

In the following, further details handling mobile device capability changes will be explained with respect to Figs. 8 and 9.

According to the present invention, for the setup of mobile device capability change information, it is suggested to compute a short description of the component and capability with respect to hardware and/or software that have changed since the left time related information was notified to the operator of a mobile device capability management device. The short description can then be communicated to the mobile device capability management device.

Heretofore, the present invention solves the problem to transfer mobile device capability change information with respect to a management tree shown in Fig. 8. The state of the management tree can change according to two ways: Nodes in the management tree can be added or removed and/or the content of a node in the management tree can change. The present invention handles both cases.

In the following, there will be described a method for retrieving changes of the contents of a leaf node in the management tree. Then will follow a method that allows to detect the deletion and insertion of nodes in the management tree. Heretofore, the following notation will used:

- bit-string - a sequence of bits with most significant bit first
- internal node - a node that is not a leaf

- c[i] denotes bit number I of bit-string c
- | denotes concatenation

In the following, there will basically be described three phases. A first phase is provided to decide what mobile device capability change information needs to be conveyed to the mobile device capability management device, which phase will be referred to as mark-up of change nodes and computation of a code word. A second phase relates to the processing of information submitted by the mobile device 10 to the mobile device capability management device 14 and the related reaction thereto, which will be referred to as system side parsing in the following. A third phase, intermediate to the first phase and the second phase, is related to the way the mobile device capability change information is exchanged from the mobile device 10 to the mobile device capability management device 14.

In the following, the first phase, i.e. mark-up of changed nodes and computation of a code word, will be explained.

Heretofore, it is assumed that in the mobile device there is maintained a management tree having a management object CHOB inserted thereto. Here, the mobile device capability change information needs to convey the changes in the management object to the mobile device capability management device 14. One option is to code this information to a bit-string c referred to as code word in the following. To compute the code word, the profile change information generating unit 22 shown in Fig. 3 may follow

different strategies, which will be explained in the following.

Firstly, the mobile device profile monitoring unit 20 shown in Fig. 3 needs to detect which changes have been made to the management object CHOB since the last time a code word c was sent to the mobile device capability management device 14, i.e. a difference between the management object CHOB at a current time and a mobile device capability systems view at the network side according to a previous time of mobile device capability management. One way of achieving this is to keep a copy or a representation of a previous management tree T' in the mobile device profile memory unit 18 shown in Fig. 3, which management tree T' was used to compute a last code word. Upon the time of computing a new code word, the mobile device profile monitoring unit 20 will compare the current management tree T with the stored representation of a previous management tree T' and mark the nodes that differ. Another approach would be that the mobile device profile monitoring unit 20 marks the nodes at the time of change, e.g., if a speech codec functionality is updated to a new version. Then, the corresponding node would be flagged as changed during the update process.

With respect to the first phase, following the mark-up of changed nodes if the profile change information generating unit 22 uses the representation of the management tree with marked change nodes and computes the code word c for transmission to the mobile device capability management device 14. The code word can be computed by the profile

change information generating unit 22 in several ways, as will be explained in the following.

Heretofore, one may assume that there has been a change in the contents of one or more nodes in the management tree, but no addition or deletion of one or more nodes in the management tree. To compute the code word, the profile change information generating unit 22 initially lets the code word be an empty string and start at the root of the management tree T, which will then be traversed in the following manner:

1. If this node is an internal node, then go to step 2.
1a. If the content of this node has changed, then c=c|1.
1.b. If the content of this node has not changed, then c=c|0.
1c. End of recursion branch.
2. For each of the children of this node execute the procedure recursively.

It should be understood that if there is no specified order enforced on the children of one node in the management tree, then step 2 referred to above might be ambiguous. To avoid this, it is suggest to specify an order on the nodes of the management tree, e.g., a lexicographical order of the unified resource identifiers URI identifying each node. The code word c resulting from the execution of the given procedures hands bit-string c[0]|c[1]|c[2]| ... |c[n-1], where n is the number of the leaves in the management tree, and c[i](0<=i<n) is equal to a value of 1 when the content of the node i has changed and 0 otherwise.

One example of the procedure outlined above is shown in Fig. 8. For this example, all shaded nodes are unchanged, but the black nodes have been marked as changed by the mark-up procedure. In part **a** of Fig. 8, the first leaf is encountered, it has not changed, so that the code word is set to 0. In part **b** of Fig. 8, the next code is encountered, it has not changed either, so a value of 0 is appended to the code word, yielding 00. In part **c** of Fig. 8, a node which is marked as changed is encountered, so that the code word is accordingly set to c=001. The recursion will then proceed to the right branch of the management tree and examine that the leaf node of this branch has not changed, part **d** and **e** of Fig. 8, so that at the end of the traversal the final code word is c=0010.

The second phase is then related to the code word set-up along the line explained above handled at the mobile communication network side. When the mobile device capability management device 14 receives the code word c from the mobile device 10, it will retrieve its current view of the management tree T' from the mobile device profile memory 36 shown in Fig. 6. Here, the change information processing unit 34 will first set a counter i to 1 traverse the management tree T' in the same order which was followed by the mobile device 10, as explained with respect to Fig. 8. Each time the change information processing unit 34 visits a leaf node of the management tree T', it looks at the corresponding value c[i] of the code word. If the related value is equal to 1, the change information processing unit 34 marks the corresponding node

as changed and increases the counter i by 1 and proceeds with the traversal. If the related value c[i] is equal to 0, a counter is increased by 1 and the traversal is continued. When the change information processing unit 34 has executed the traversal of the tree and has looked at all bits of the code word, the change information processing unit may then trigger a command to read the corresponding nodes marked as changed from the management tree maintained at the mobile device 10.

In the following, further modifications of the first and second phase, i.e. marking of changed nodes at the mobile device 10 and computation of a code word, followed by system side parsing will be explained. Instead of sending the code word c to the mobile device capability management device 14, the mobile device 10 may choose to send a hash of the code word c to keep the mobile device capability change message short and/or of fixed length. Then, the mobile device capability management device 14 would keep a pre-computed list of hashes in the mobile device profile memory 36, and further a one-to-one mapping between which nodes have changed and the set of hashes that would correspond thereto. Here, the hash value should be shorter than the code word, which might give raise to collisions, i.e. if there are several possible node-change configurations corresponding to a single hash value. However, there might be several of the configurations that are not possible, e.g., some nodes may not be changed in the management tree or only be changed by the operator of the mobile communication network, so that this problem may be easily overcome. Optionally, for long code words one may

also apply a compression scheme to reduce amount of traffic in the mobile communication network.

While above it has been explained that the code word c contains one bit per leaf in the management tree, it may also be possible to reduce the length of the code word by skipping bits of those nodes that are known to have a static content. E.g., on most mobile devices it may not be possible to upgrade the screen, and in that case, the scree size of pixels could be skipped as an example. The pre-requisite is that both the mobile device 10 and the operator of the mobile communication network take the same viewpoint that the nodes are static.

Further, while above only tree traversal algorithms have been given as example, in the most general manner it is important that the management data structure is traversed in the same order by the mobile device 10 and the mobile device capability management device 14.

In the following, with respect to Fig. 9, there will be given another example for generation of mobile device capability change information with respect to the detection of additional deletion of nodes. Heretofore, a slightly different code c can be constructed.

As shown in Fig. 9, instead of only assigning a bit to each leaf node of the management tree, the profile change information generating unit 22 now assigns a bit to each node in the management tree. The value of the bit is assigned as follows: If the node has no new children or

does not have any children removed, the value of the bit associated with the node is 0, otherwise it is 1. The assignment of the bits is performed during a marking process as described above.

Heretofore, the mobile device profile monitoring unit 20 traverses the management tree using, e.g., a breadth first search, and output the bit assigned to each node when the node is visited. Also, mobile change information generating unit 22 must not traverse any new sub-trees, but only mark the parent new sub-trees with a value of 1.

Fig. 9 shows an example of the approach so far. Here, the code word computed using a breadth first search for the management tree would lead to a code word c=00010000. The order of the nodes using the breadth first search is generated by the profile change information generating unit 22 reading the nodes in the management trees top-down, left-to-right.

Then, when the mobile device capability management device 14 receives the code word, it will be scanned by the exchange information processing unit 34 from left to right, following the management tree, as shown in Fig. 9. Here, the change information processing unit 34 should use the same tree traversal algorithm as it was used by the profile change information generating unit 22 in the mobile device. If the change information processing unit 34 processes a node of which the corresponding bit in the code word is 1, this will be marked for further investigation. Once the change information processing unit 34 has executed the

scanning of the code word c, it then issues a command to the mobile device 10 for traversal of the marked nodes and pick-up of any changes that have occurred to the related children nodes. While it would of course be possible to read from the mobile device 10 during a scan, it is important to not include the newly found or to remove the deleted nodes from the management tree before the scan is complete.

The advantage of the approach explained above is that it is not necessary to read all nodes of the management tree in the mobile device 10 for upgrade of a management tree in the mobile device capability management device 14. Only sub-trees that are new or have been deleted need to be visited.

In the following, the third phase being related to exchange of mobile capability change information between the mobile device and the mobile device capability management device 14 will be explained with respect to Fig. 10.

As shown in Fig. 10, the mobile device capability change information and/or related change message CHAN can be communicated using traditional cellular signaling S87, such as automatic device detection MSC-Home Location Register HLR-Device Management Server. Depending on the bandwidth available the granularity of the mobile device profile change information may be adapted.

As shown in Fig. 10, it is also possible to convey such mobile device capability information upon GPRS attachment

using a RADIUS, 'Remote Authentication Dial In User Service', RFC2138, IETF, April 1997 / DIAMETER, 'Diameter Base Protocol', RFC3588, IRTF, September 2003 message as an attribute value from the basic GPRS support node GGSN to the AAA unit and then further to the device management GMS. Further, the mobile device can convey the mobile device capability change information via the service layer, e.g., upon start of an open mobile alliance device management session according to MTREE, OMA, SyncML Device Management Tree and Description, version 1.1, February 2002.

While with respect to Fig. 10 the application of a cellular network signaling of a GPRS/PS network and/or the service layer has been described, these processes are to be considered as examples only, and any other suitable approach to information exchange may be applied within the framework of the present invention.

CLAIMS

1. Method of operating a mobile device capability management device (14) in a mobile communication network with respect to at least one mobile device (10) served by the mobile communication network, wherein a mobile device capability is represented as first a mobile device profile in the mobile device capability management device and as a second mobile device profile in the mobile device, comprising the steps:

   - receiving (S16) a mobile device capability change information with respect to a mobile device capability change;

   - assessing (S18) the need to update the first mobile device profile on the basis of a mobile device capability management policy and the received mobile device capability change information.

2. Method according to claim 1, ***characterized in that*** it comprises a step of updating (S36) the first mobile device profile according to the mobile device capability management policy and the mobile device capability change information.

3. Method according to claim 2, ***characterized in that*** the step of updating the first mobile device profile comprises the steps:

- establishing (S32) a mobile device management session between the mobile device capability management device and the mobile device for evaluation of the second mobile device profile with respect to a mobile device capability change; and

- receiving (S34) the mobile device capability update information from the mobile device in accordance with an evaluation result.

4. Method according to one of the claim 1 to 3, ***characterized in that*** the first mobile device profile and/or the second mobile device profile is represented as a mobile device management tree structure comprising mobile device management objects.

5. Method according to claim 4, ***characterized in that*** the step of updating the first mobile device profile in the mobile device capability management device comprises a step of updating a mobile device capability change object (CHOB), wherein the mobile device capability change object is a mobile device management object in the mobile device management tree structure dedicated to mobile device capability change management.

6. Method according to one of the claims 1 to 5, ***characterized in that*** the mobile device capability change information (CHAN) contains a qualitative parameter indicating seriousness of change to the second mobile device profile and a quantitative parameter indicating amount of effective changes to the second mobile device profile.

7. Method according to claim 6, ***characterized in that*** the step of assessing (S18) the need to update the second mobile device profile is achieved through classification of the quantitative parameter and/or the qualitative parameter.

8. Method according to one of the claims 1 to 7, ***characterized in that*** it comprises a step of notifying (S38) a service provider providing at least one service to the mobile device and/or an application interacting with the mobile device on a mobile device capability change.

9. Method according to one of the claims 1 to 8, ***characterized in that*** mobile device capability change information is received directly from the mobile device at the mobile device capability management device (14).

10. Method according to one of the claims 1 to 8, ***characterized in that*** mobile device capability change information is received from a mobile device capability monitoring device.

11. Method according to one of the claims 1 to 10, ***characterized in that*** the mobile device capability change information is exchanged through circuit switched mobile communication, packet switched mobile communication, and/or through service layer data exchange.

12. Method of operating a mobile device (10) being enabled for mobile device capability management through a mobile communication network, wherein a mobile device

capability state is represented in the mobile device as a mobile device profile, comprising the steps:

- monitoring (S10) a change in the mobile device capability;

- generating (S12) mobile device capability change information with respect to a mobile device capability change;

- forwarding (S14) the mobile device capability change information to the mobile communication network.

13. Method according to claims 12, ***characterized in that*** the mobile device capability change information comprises a qualitative parameter indicating seriousness of change to the mobile device profile and a quantitative parameter indicating amount of effective changes to the mobile device profile.

14. Method according to claim 12 or 13, ***characterized in that*** it comprises the steps:

- receiving (S22) a request for detailed mobile device capability change information from the mobile communication network;

- establishing (S24) a mobile device management session with the mobile communication network; and

- forwarding (S26) the detailed mobile device capability change information to the mobile communication network.

15. Method according to one of the claims 12 to 14, ***characterized in that*** it comprises a step of representing the mobile device profile as a mobile device management tree structure comprising at least one mobile device management object.

16. Method according to claim 15, ***characterized in that*** it comprises a step of representing the mobile device capability change information as a mobile device capability change object (CHOB) in the mobile device management tree, wherein the mobile device capability change object (CHOB) is a mobile device management object in the mobile device management tree structure dedicated to mobile device capability change management.

17. Method according to one of the claims 12 to 16, ***characterized in that*** the mobile device capability change information is forwarded to a mobile device capability management device (14) in the mobile communication network.

18. Method according to one of the claim 12 to 16, ***characterized in that*** the mobile device capability change information is forwarded to a mobile device capability monitoring device in the mobile communication network, for subsequent conveyance thereof to a mobile device capability management device (14) in the mobile communication network.

19. Method according to one of the claims 12 to 19, ***characterized in that*** the mobile device capability change information is exchanged with the mobile communication network through circuit switched mobile

communication, packet switched mobile communication, and/or through service layer data exchange.

20. Mobile device capability management device (14) being operated in a mobile communication network with respect to at least one mobile device (10) served by the mobile communication network, wherein a mobile device capability is represented as a first mobile device profile in the mobile device capability management device and as a second mobile device profile in the mobile device, comprising:

- a communication unit (26) adapted to receive mobile device capability change information with respect to a mobile device capability change;

- an assessment unit (28) adapted to assess the need to update the first mobile device profile on the basis of a mobile device capability management policy and the mobile device capability change information.

21. Mobile device capability management device according to claim 20, ***characterized in that*** it comprises a mobile device profile updating unit (30) adapted to update the first mobile device profile according to the mobile device capability management policy and the received mobile device capability change information.

22. Mobile device capability management device according to claim 21, ***characterized in that*** the mobile device profile updating unit (30) comprises:

- a session initiating unit (32) adapted to initiate a mobile device management session

between the mobile device capability management device and the mobile device for evaluation of the second mobile device profile with respect to a mobile device capability change; and

- a capability change information processing unit (34) adapted to process the mobile device capability change information from the mobile device in accordance with an evaluation result.

23. Mobile device capability management device according to one of the claim 20 to 22, ***characterized in that*** it comprises a mobile device profile memory (36) adapted to store the first mobile device profile and/or the second mobile device profile according to a mobile device management tree structure comprising mobile device management objects.

24. Mobile device capability management device according to claim 23, ***characterized in that*** the capability change information processing unit (34) is adapted to update the first mobile device profile in the mobile device profile memory by updating a mobile device capability change object, wherein the mobile device capability change object (CHOB) is a mobile device management object in the mobile device management tree structure dedicated to mobile device capability change management.

25. Mobile device capability management device according to one of the claims 20 to 24, ***characterized in that*** the assessment unit (28) is adapted to process mobile device capability change information containing a qualitative parameter indicating seriousness of change to the second mobile device profile and a quantitative

parameter indicating amount of effective changes to the second mobile device profile.

26. Mobile device capability management device according to claim 25, ***characterized in that*** the assessment unit (28) is adapted to assess the need to update the second mobile device profile through classification of the quantitative parameter and/or the qualitative parameter.

27. Mobile device capability management device according to one of the claims 20 to 26, ***characterized in that*** the mobile device profile updating unit (30) is adapted to notify a service provider (16) providing at least one service to the mobile device and/or an application interacting with the mobile device on a mobile device capability change.

28. Mobile device capability management device according to one of the claims 20 to 27, ***characterized in that*** communication unit (26) is adapted to receive the mobile device capability change information directly from the mobile device.

29. Mobile device capability management device according to one of the claims 20 to 27, ***characterized in that*** communication unit (26) is adapted to receive the mobile device capability change information from a mobile device capability monitoring device.

30. Mobile device capability management device according to one of the claims 20 to 29, ***characterized in that*** the communication unit (26) is adapted to receive the mobile device capability change information through circuit switched mobile communication, packet switched

mobile communication, and/or through service layer data exchange.

31. Mobile device (10) being enabled for mobile device capability management through a mobile communication network, wherein a mobile device capability state is represented in the mobile device as a mobile device profile, comprising:

- a mobile device profile monitoring unit (20) adapted to monitor a change in the mobile device capability;

- a profile change information generating unit (22) adapted to generate mobile device capability change information with respect to a mobile device capability change;

- a communication unit (24) adapted to forward the mobile device capability change information to the mobile communication network.

32. Mobile device according to claims 31, ***characterized in that*** the profile change information generating unit (22) is adapted to generate the mobile device capability change information with a qualitative parameter indicating seriousness of change to the mobile device profile and a quantitative parameter indicating amount of effective changes to the mobile device profile.

33. Mobile device according to claim 31 or 32, ***characterized in that*** the communication unit (24) is adapted to:

- receive a request for update of a remote mobile device profile being maintained in the mobile communication network;

- establish a mobile device management session with the mobile communication network; and to

- forward detailed mobile device capability change information to the mobile communication network.

34. Mobile device according to one of the claims 31 to 33, ***characterized in that*** it comprises a mobile device profile memory unit (18) adapted to store the mobile device profile as a mobile device management tree structure comprising at least one mobile device management object.

35. Mobile device according to claim 34, ***characterized in that*** the profile change information generating unit (22) is adapted to generate mobile device capability change information as a mobile device capability change object in the mobile device management tree, wherein the mobile device capability change object is a mobile device management object in the mobile device management tree structure dedicated to mobile device capability change management.

36. Mobile device according to one of the claims 31 to 35, ***characterized in that*** the communication unit (24) is adapted to forward the mobile device capability change information to a mobile device capability management device in the mobile communication network.

37. Mobile device according to one of the claim 31 to 35, ***characterized in that*** the communication unit (24) is

adapted to forward the mobile device capability change information to a mobile device capability monitoring device in the mobile communication network, for subsequent conveyance thereof to a mobile device capability management device (14) in the mobile communication network.

38. Mobile device according to one of the claims 31 to 37, ***characterized in that*** the communication unit (26) is adapted to exchange mobile device capability change information with the mobile communication network through circuit switched mobile communication, packet switched mobile communication, and/or through service layer data exchange.

39. Computer program product directly loadable into the internal memory of a mobile device capability management device (14) operated in a mobile communication network, comprising software code portions for performing the steps of one of the claims 1 to 11, when the product is run on a processor of the mobile device capability management device.

40. Computer program product directly loadable into the internal memory of a mobile device (10) being enabled for mobile device capability management through a mobile communication network, comprising software code portions for performing the steps of one of the claims 12 to 19, when the product is run on a processor of the mobile device.

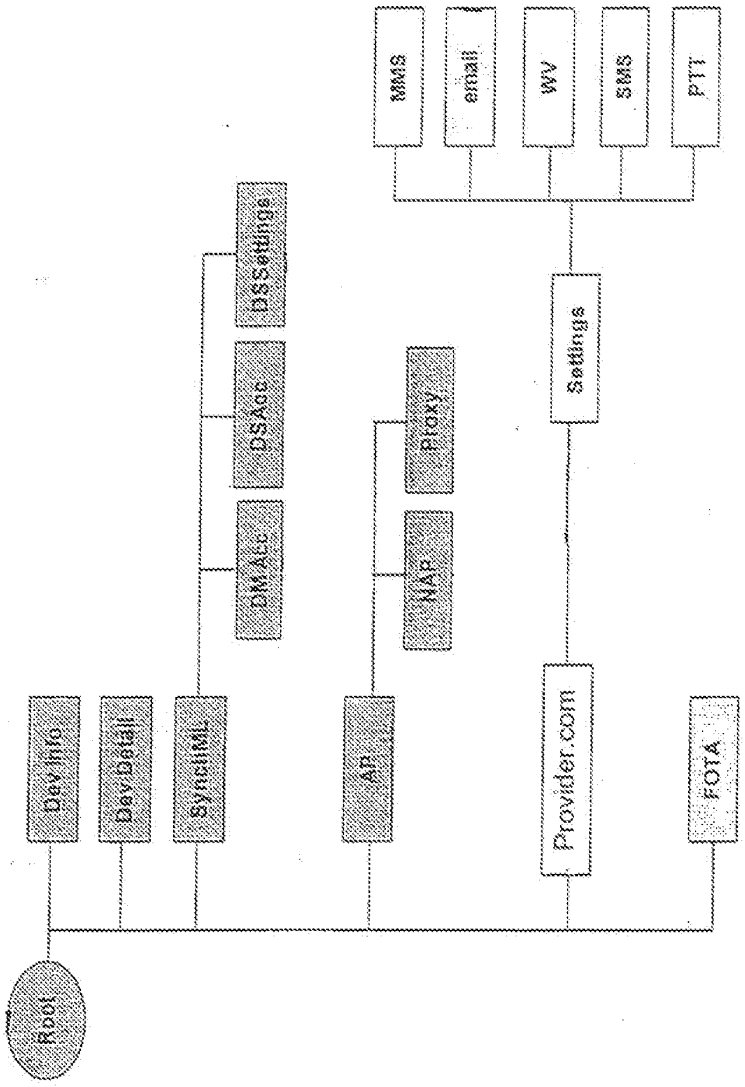

MMS
email
WV
SMS
PTT
DSSetting
DSAcc
DMAcc
Proxy
NAP
Settings
DevInfo
DevDetail
SyncML
AP
Provider.com
FOTA
Root

Fig. 1

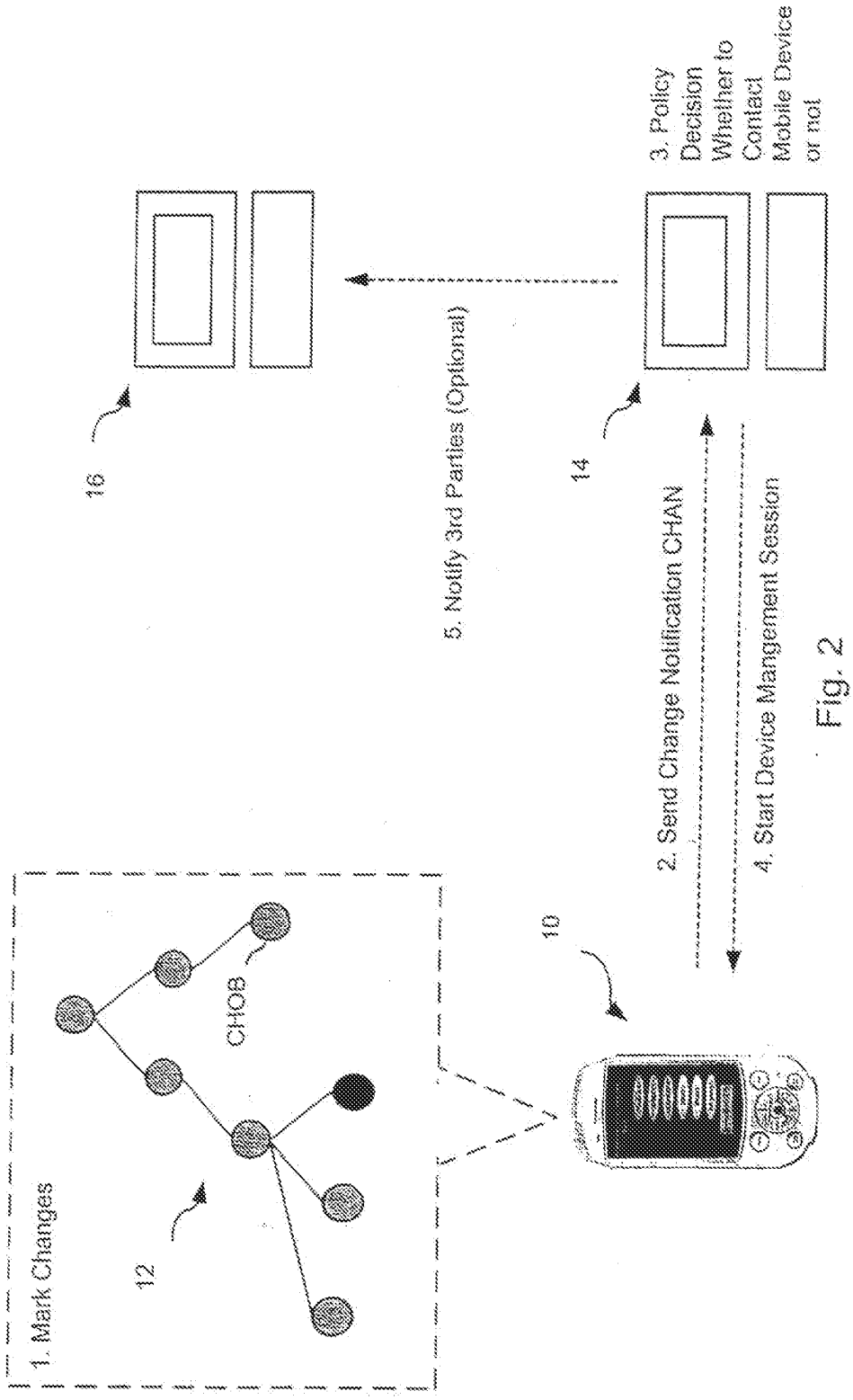

1. Mark Changes
12
CHOB
10
2. Send Change Notification CHAN
4. Start Device Mangement Session
14
3. Policy Decision Whether to Contact Mobile Device or not
5. Notify 3rd Parties (Optional)
16

Fig. 2

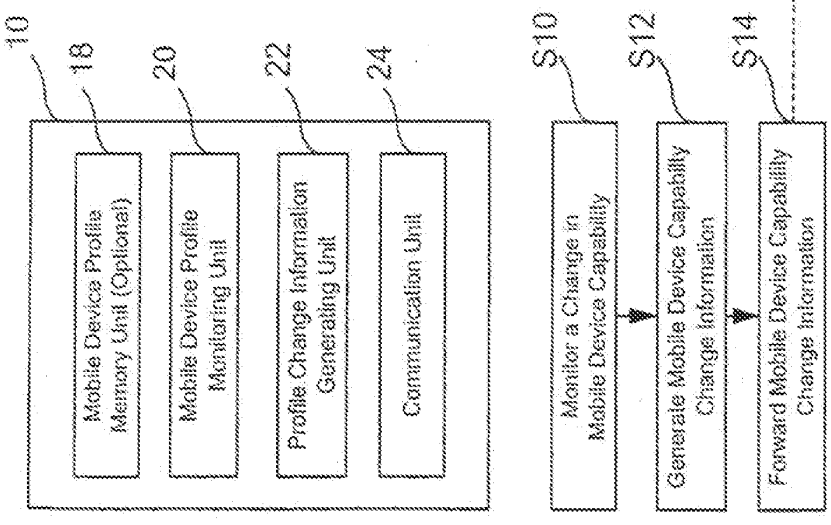

10
18
20
22
24
Mobile Device Profile Memory Unit (Optional)
Mobile Device Profile Monitoring Unit
Profile Change Information Generating Unit
Communication Unit
S10
S12
S14
Monitor a Change in Mobile Device Capability
Generate Mobile Device Capability Change Information
Forward Mobile Device Capability Change Information

Fig. 3

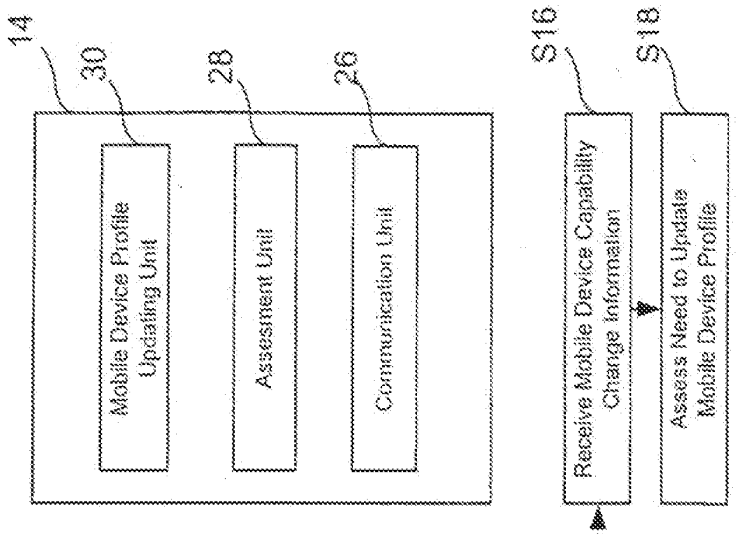

14
30
28
26
Mobile Device Profile Updating Unit
Assessment Unit
Communication Unit
S16
S18
Receive Mobile Device Capability Change Information
Assess Need to Update Mobile Device Profile

Fig. 4

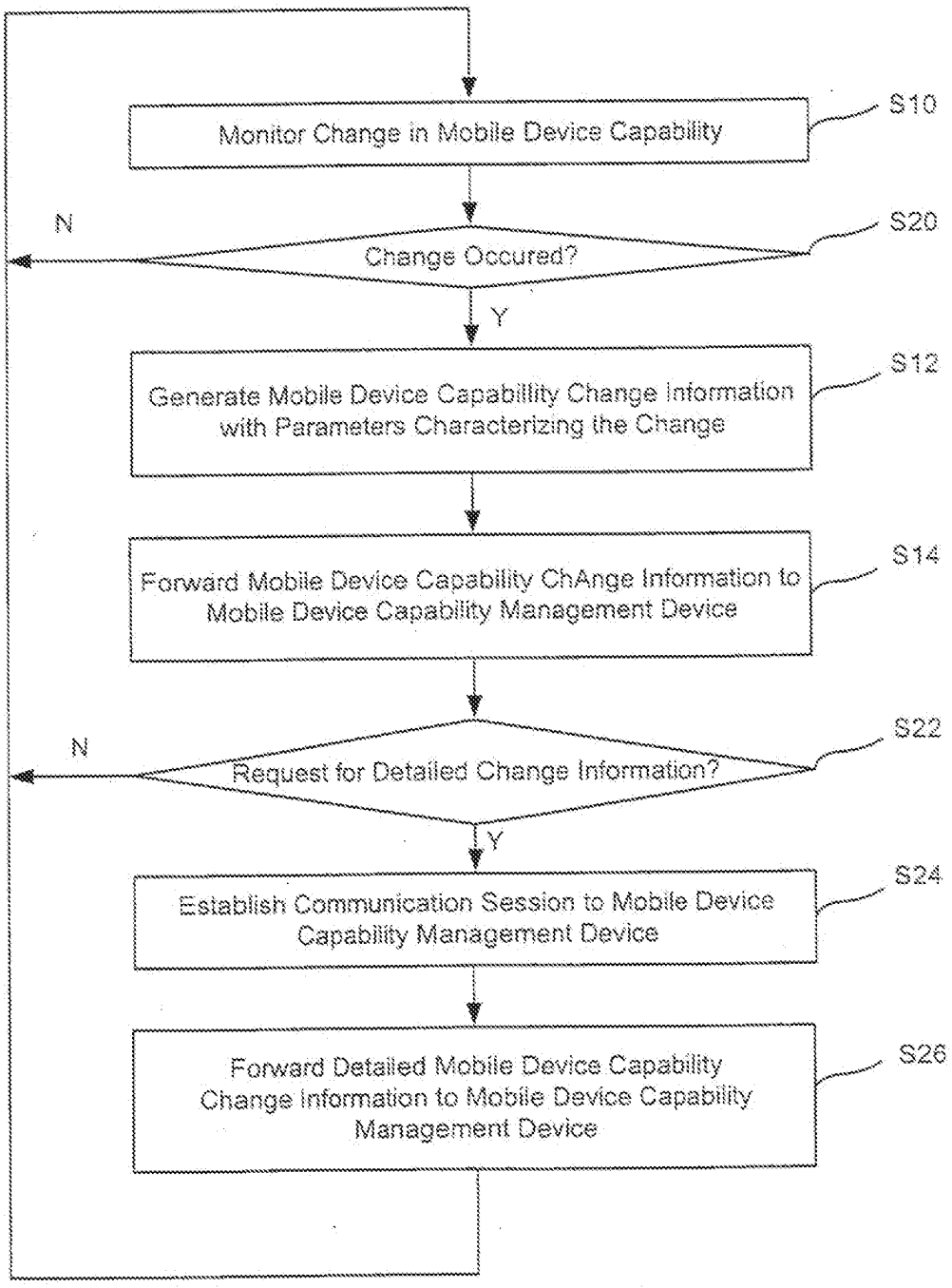

Monitor Change in Mobile Device Capability
S10
N
Change Occured?
S20
Y
Generate Mobile Device Capabillity Change Information with Parameters Characterizing the Change
S12
Forward Mobile Device Capability ChAnge Information to Mobile Device Capability Management Device
S14
N
Request for Detailed Change Information?
S22
Y
Establish Communication Session to Mobile Device Capability Management Device
S24
Forward Detailed Mobile Device Capability Change Information to Mobile Device Capability Management Device
S26

Fig. 5

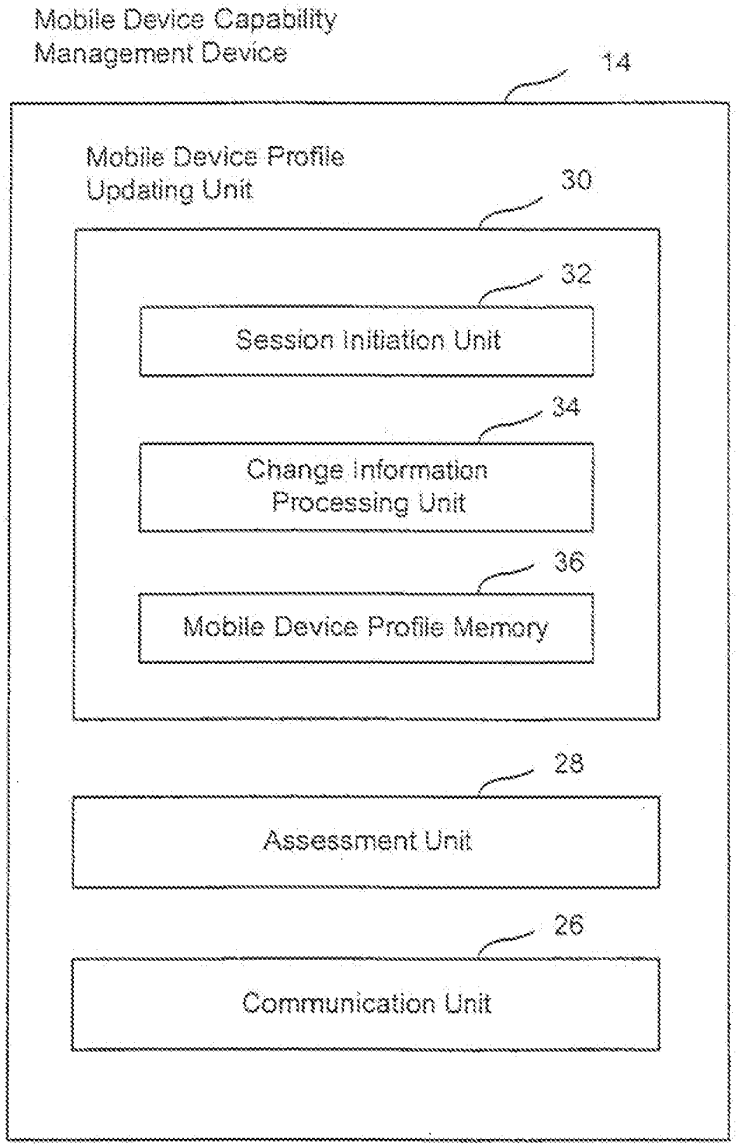

Mobile Device Capability Management Device
14
Mobile Device Profile Updating Unit
30
32
Session Initiation Unit
34
Change Information Processing Unit
36
Mobile Device Profile Memory
28
Assessment Unit
26
Communication Unit

Fig. 6

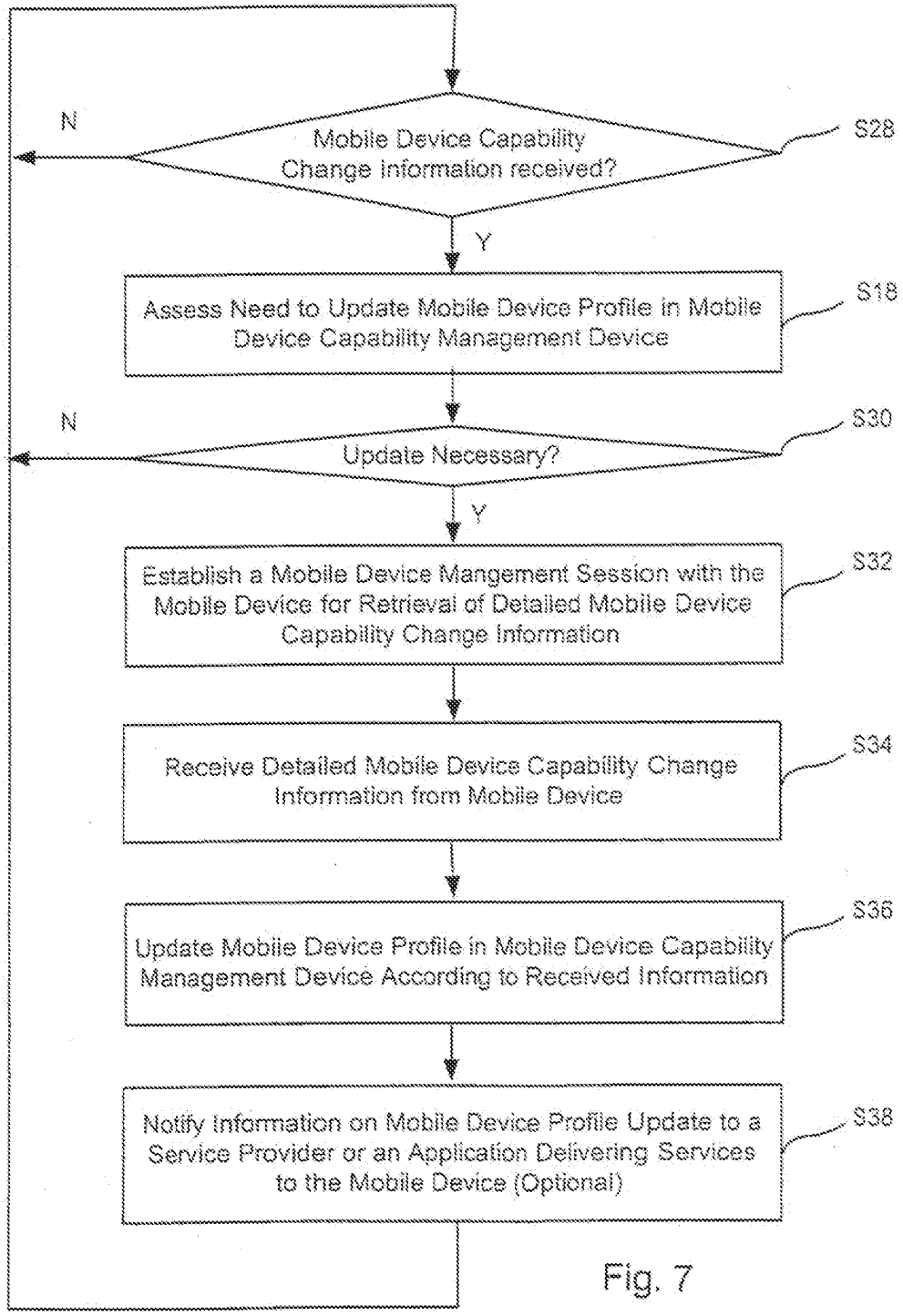

Mobile Device Capability Change Information received?
S28
N
Y
Assess Need to Update Mobile Device Profile in Mobile Device Capability Management Device
S18
Update Necessary?
S30
N
Y
Establish a Mobile Device Mangement Session with the Mobile Device for Retrieval of Detailed Mobile Device Capability Change Information
S32
Receive Detailed Mobile Device Capability Change Information from Mobile Device
S34
Update Mobile Device Profile in Mobile Device Capability Management Device According to Received Information
S36
Notify Information on Mobile Device Profile Update to a Service Provider or an Application Delivering Services to the Mobile Device (Optional)
S38

Fig. 7

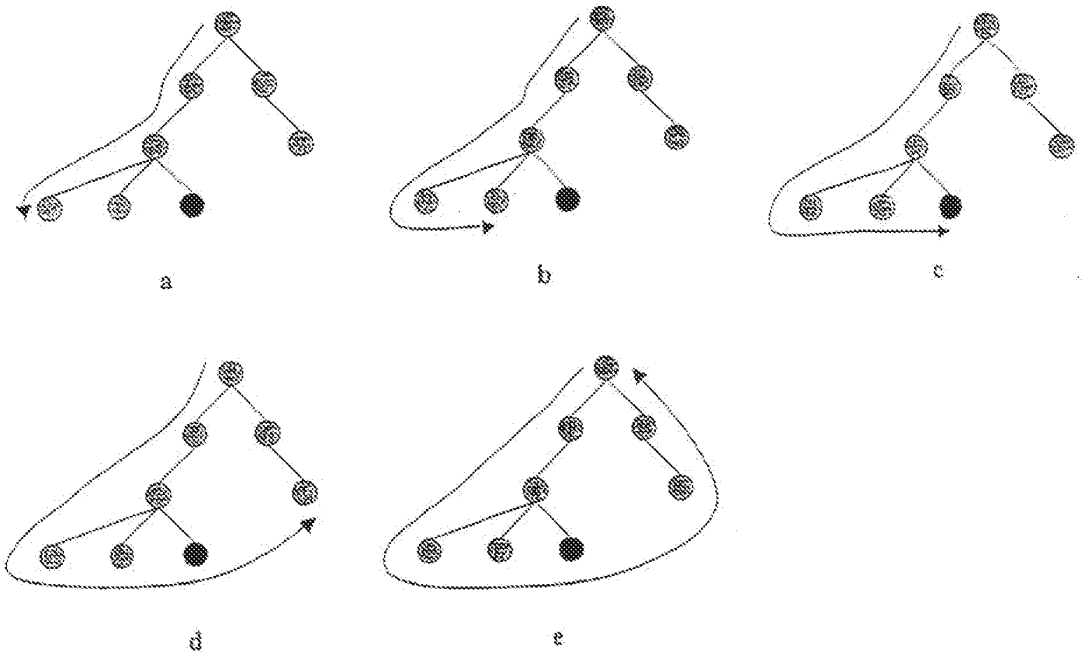


Fig. 8

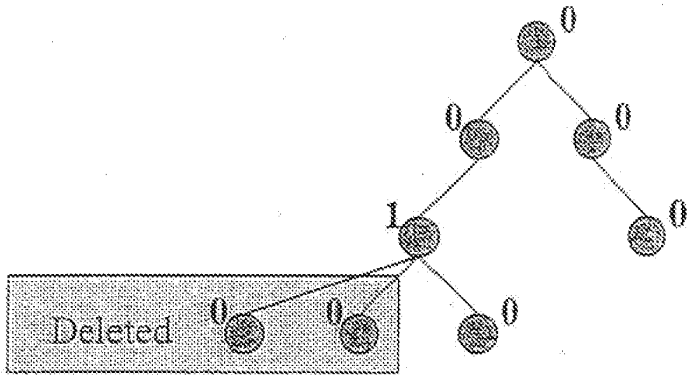


Fig. 9

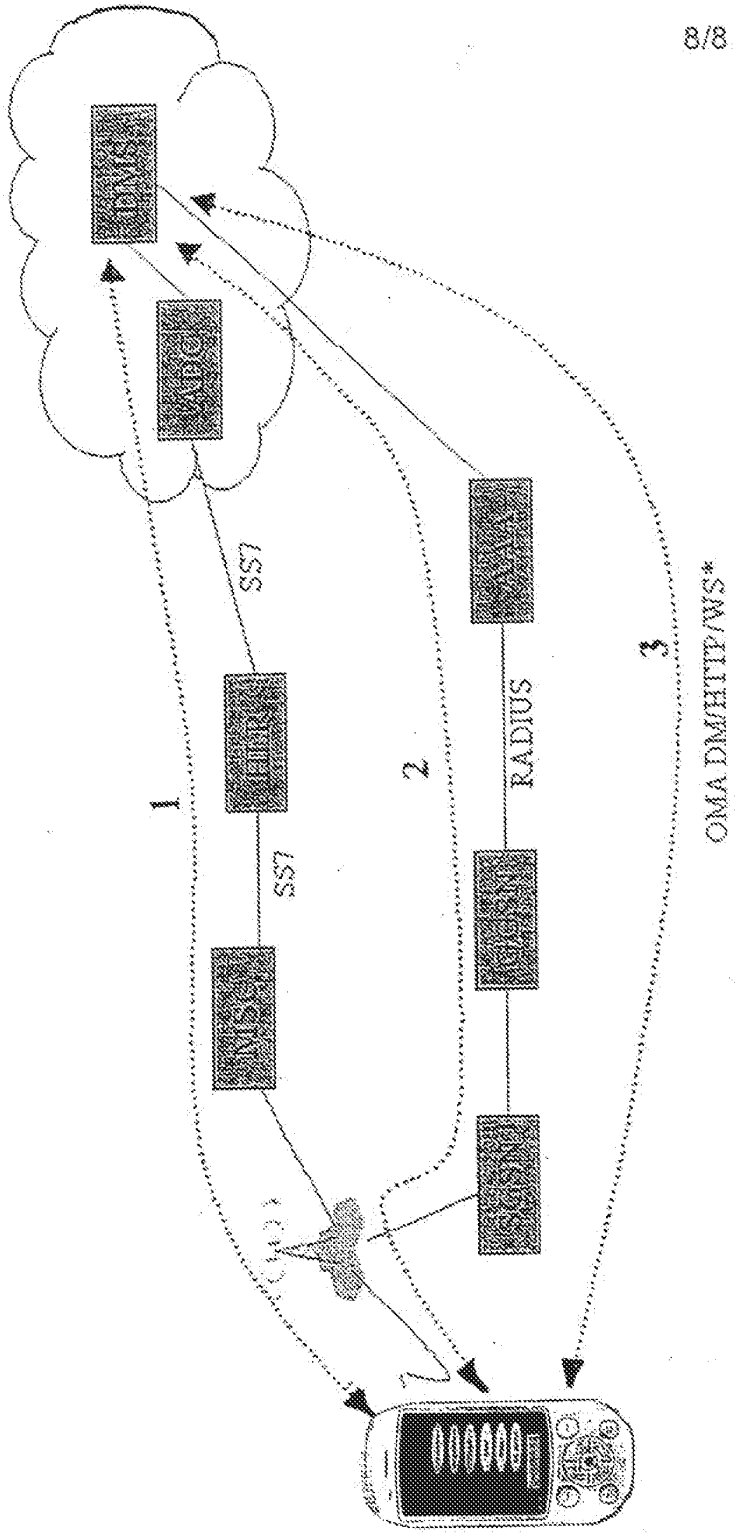


Fig. 10

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/006776

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04Q7/38 H04L29/08 H04M3/42

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04Q H04L H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1 376 989 A (OPENWAVE SYSTEMS INC) 2 January 2004 (2004-01-02) abstract paragraphs [0060], [0069] - [0071]	1-40
Y	WO 02/091702 A (SIEMENS AKTIENGESELLSCHAFT; KUNZ, ALBRECHT; LEITGEB, MANFRED; SWETINA,) 14 November 2002 (2002-11-14) abstract page 6, line 18 - line 24 page 8, line 5 - line 25 page 9, line 21 - line 27 page 11, line 20 - line 35	1-3,7-9, 11-14, 17, 19-22, 26-28, 30-33, 36,38-40

-/--

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 October 2006	17/10/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Rüschmann, Frank

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/006776

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DEVICE MANAGEMENT WORKING GROUP: "OMA Device Management Tree and Description OMA-TS-DM-TND-V1_2-20050615-C" OMA, 15 June 2005 (2005-06-15), pages 1-48, XP002360958 page 4, paragraph 4 page 5, paragraph 5 page 23, paragraph 8 - page 24, paragraph 8.2 page 25, paragraphs 9,9.1	4-6,10, 15,16, 18, 23-25, 29,34, 35,37, 39,40
A	EP 0 685 972 A (NOKIA MOBILE PHONES LTD; NOKIA CORPORATION) 6 December 1995 (1995-12-06) column 1, line 10 - line 27 column 8, line 7 - line 36 column 9, line 1 - line 36	1,8,9, 11,17, 19,20, 27,28, 30,36, 38-40
A	WO 00/67501 A (NOKIA MOBILE PHONES LTD; TERVO, VESA; MAEKELAE, SATU) 9 November 2000 (2000-11-09) page 13, line 3 - page 14, line 16 page 16, line 32 - page 17, line 34 page 26, line 21 - page 27, line 14	1,8,9, 11,17, 19,20, 27,28, 30,36, 38-40

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2006/006776

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1376989 A	02-01-2004	US 2004024867 A1	05-02-2004
WO 02091702 A	14-11-2002	EP 1386466 A1	04-02-2004
EP 0685972 A	06-12-1995	DE 69532597 D1	01-04-2004
		DE 69532597 T2	16-12-2004
		FI 942548 A	01-12-1995
WO 0067501 A	09-11-2000	AU 4407100 A	17-11-2000
		CN 1359599 A	17-07-2002
		EP 1175797 A1	30-01-2002
		FI 991000 A	31-10-2000
		JP 2002543720 T	17-12-2002

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
12 January 2006 (12.01.2006)

(10) International Publication Number
WO 2006/004467 A1

(51) International Patent Classification[7]: H04Q 7/38, H04L 12/56

(21) International Application Number: PCT/SE2004/001087

(22) International Filing Date: 5 July 2004 (05.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **TELEFONAKTIEBOLAGET LM ERICSSON (publ)** [SE/SE]; S-164 83 Stockholm (SE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SKOG, Robert** [SE/SE]; Gullvivegränd 7, S-165 76 Hässelby (SE). **PETERSSON, Justus** [SE/SE]; Sankt Paulsgatan 28B, S-118 48 Stockholm (SE).

(74) Agent: **DR LUDWIG BRANN PATENTBYRÅ AB**; P.O. Box 171 92, S-104 62 Stockholm (SE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

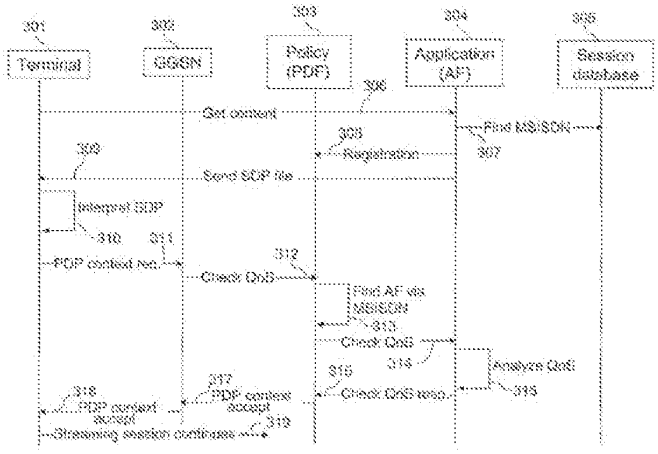
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: BINDING MECHANISM FOR QUALITY OF SERVICE MANAGEMENT IN A COMMUNICATION NETWORK



(57) Abstract: The present invention relates to methods and arrangements for setting up a policy controlled bearer service for delivery of a service from an application function (304) to a mobile terminal (301). According to the present invention a terminal identity of the mobile terminal may be used to bind to applicable policy information when setting up the bearer service. Thus the present invention obviates the need to transmit specific binding information to the mobile terminal during application layer signaling for the purpose of policy controlled bearer service set-up. The terminal identity is registered (308) in a register of a policy decision function (303) along with policy identifying information. When a bearer service activation request is received from the mobile terminal (301) the request is mapped (313) to applicable policy information by means of the terminal identity and the matching registration in the register of the policy decision function (303).

WO 2006/004467 A1

BINDING MECHANISM FOR QUALITY OF SERVICE MANAGEMENT IN A COMMUNICATION NETWORK

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is related to commonly-assigned co-pending international patent application No. PCT/SE04/_____, entitled "Devices and Methods for Push Message Initiated Service", filed on July 5, 2004; international patent application No. PCT/SE04/_____, entitled "Methods and Devices for Supplying Quality of Service Parameters in HTTP Message", filed on July 5, 2004; and international patent application No. PCT/SE04/_____, entitled "Methods and Devices for Changing Quality of Service", filed on July 5, 2004, the disclosures of which are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to equipment and methods in mobile communication systems, and more particularly, to mechanisms for quality of service management of transmissions relating to a packet data service.

BACKGROUND

Communication networks for packet based communication of information in the form of data bits are well known to the person skilled in the art. The growing importance of mobile communication creates the demand to transfer data over wireless connections. The 3rd Generation Partnership Project (3GPP) is a collaboration agreement that brings together a number of telecommunications standards bodies and was established to produce globally applicable Technical Specifications and Technical Reports for a 3rd Generation Mobile System based on evolved GSM core networks and the radio access technologies that they support. This includes maintenance and development of the Global System for Mobile communication (GSM) Technical Specifications and Technical Reports including evolved radio access technologies (e.g. General Packet Radio Service (GPRS) and Enhanced Data rates for GSM Evolution (EDGE)).

One of the issues that the 3GPP is currently working on is a framework for providing end-to-end Quality of Service for services involving GPRS networks as described in the technical specification 3GPP TS 23.207 V6.2.0 (2004-03). This technical specification describes the use of a Policy Control mechanism in order to control and authorize the use of network resources prior to set-up of a service.

In the future it is assumed that packet switched services undergo Policy Control by means of a Policy Decision Function (PDF). The PDF is a logical element that is used as a logical policy decision element and enables coordination between events in the application layer and resource management in the IP bearer layer.

Policy Control was developed since it is considered important that network managers and service providers can to monitor, control, and enforce the use of network resources and services based on "policies" derived from certain criteria such as the identity/authority level of users and applications, traffic bandwidth requirements, security considerations, time of day/week etc. Because there are varying circumstances in which various entities are entitled to use the services they request, there is a need for rules, need for enforcement methods of these rules, and a need for a "judge" to decide when they apply. Accordingly, three major components of a policy system include policy rules, which are typically stored in a policy database, policy enforcement, which may be implemented at Policy Enforcement Points (PEP), and Policy Decision Points. The IETF has standardized a protocol for information exchange between PEPs and Policy Decision Points under the term Common Open Policy Service (COPS). In general, a policy may be regarded as a collection of rules that result in one or more actions when specific conditions exist.

Fig. 1 is a schematic block diagram that illustrates the different elements that are involved in the policy control mechanism developed by 3GPP. Fig. 1 illustrates an example where a service is set-up between a mobile User Equipment (UE) 1 and an Application Function (AF) of a service provider. The

UE includes an AF Client 3 which communicates with the AF 2 by means of AF session signaling 4 in order to initiate an application session. The set-up of the session and activation of the service require set-up of a bearer service to carry the data packets of the session. The session may include one or a plurality of different IP flows with specific requirements for bearer resources. The AF session signaling 4 may include information on Quality of Service (QoS) requirements which are translated into bearer QoS parameters by a translation/mapping function 5 in the UE. In order to activate the bearer service the UE sends a request for activation of a PDP context with the bearer QoS parameters that are determined by the translation/mapping function to a GGSN 6. The communication between the UE and the GGSN is in this example handled by means of UMTS BS Managers 7, 8 in the UE and the GGSN respectively. In this example the GGSN functions as a PEP and is provided with an IP BS Manager 9 which is involved in QoS management. The GGSN 6 interacts with a PDF 10, via a Go interface 11, in order to execute policy control. The PDF 10 makes policy decisions based on policy set-up information obtained from the AF 2 via a Gq interface 12. More detailed information regarding the elements illustrated in Fig. 1 and their function can be found in the above mentioned technical specification TS 23.207 V6.2.0.

When 3GPP specified the policy control mechanism an authorization token was also introduced in order to establish a binding mechanism for mapping PDP contexts and IP flows to policy information. The authorization token is generated by the PDF for a particular AF session on request from the AF. The authorization token includes a fully qualified domain name of the PDF and is unique across all PDP contexts associated with an APN (Access Point Name). The AF sends the authorization token to the UE in the AF session signalling. The UE should then include the applicable authorization token in any request to activate or modify a PDP context. The authorization token will then be used in the GGSN and PDF to map to the correct policy information.

The plan to introduce the authorization token and the associated handling of it when implementing policy control in 3G systems has a number of

disadvantages. One major disadvantage is that it requires modification of existing UEs in order to make them capable to handle authorization tokens. Another disadvantage is that the introduction of the authorization token requires new standardization for the RTSP (Real-Time Streaming Protocol), SIP (Session Initiation Protocol), and SDP (Session Description Protocol) and other protocols used in set-up and handling of application sessions.

SUMMARY OF THE INVENTION

An object of the present invention is to provide devices and methods for employing a simple binding mechanism when setting up a policy controlled bearer service for delivery of a service to a mobile terminal.

The above stated object is achieved by means of methods according to claim 1 and 10, and by means of a policy decision function according to claim 19.

The basic idea of the present invention is to use a terminal identity as binding information in order to bind a bearer service activation request to the applicable policy information and thereby obviate the need to transmit an authorization token to the terminal during bearer service activation.

According to a first aspect of the present invention a method for setting up a policy controlled bearer service for delivery of a service from an application function to a mobile terminal is provided. The method includes the steps of registering a terminal identity of the terminal associated with policy identifying information in a register of a policy decision function; receiving a bearer service activation request from the terminal, via a core network node, which bearer service activation request includes the terminal identity, such as MSISDN, IMEI , IMSI etc. and a set of requested QoS parameters for the bearer service; binding the bearer service activation request to the policy identifying information by means of the terminal identity to determine applicable policy information for the bearer service; performing a policy check by means of checking if the requested QoS parameters fulfill said applicable policy information; and sending an accept message to the core network node