A Multi-Agent Systems Approach for Fraud Detection in Personal Communication Systems

Suhayya Abu-Hakima Mansour Toloo Tony White
Seamless Personal Information Networking Group
Institute for Information Technology
National Research Council of Canada
Ottawa, Canada K1A 0R6
abu-hakima@iit.nrc.ca or suhayya@ai.iit.nrc.ca
tel: 613-991-1231 fax: 613-952-7151 web: www.nrc.ca/iit

From: AAAI Technical Report WS-97-07. Compilation copyright © 1997, AAAI (www.aaai.org). All rights reserved.

Abstract

The fraudulent use of telecommunications networks is costing the industry billions of dollars in lost revenue each year. The opportunities for fraudulent activity are particularly great in mobile communications as a result of the transport medium and the desire of network operators to attract large numbers of users quickly. This paper reports on the problems currently faced by mobile phone network operators and the technological solutions currently being used or implemented. A multi-agent systems approach which is expected to provide greater user customization of handset facilities and enhanced real time fraud detection is proposed.

1. Introduction: Why is fraud a key issue in mobile communication networks?

In exploring why fraud is an issue in mobile communication networks, it is essential to ask why criminals are attracted to cellular phones? In responding to this question, one has to think like a criminal. Criminals are attempting to carry out illicit activities without being traced. Mobile communication networks are based on radio communications which serve as ideal communication channels for criminals since a call is very difficult if not impossible to trace on a radio link. In addition, cellular phone service is easily stolen from legitimate subscribers. Criminals make use of mobile communication scanners to easily clone a Mobile Identification Number (MIN) and its associated Equipment Serial Number (ESN). MIN/ESN pairs are normally broadcast from an active user's cellular phone so that it may legitimately communicate with the mobile network. Criminals cycle through stolen MIN/ESN pairs making such fraud very difficult to detect and trace. Mobile networks provide unlimited international roaming which is ideal for criminal activity. Mobile networks are as perfect for drug dealers, bookies and criminals as they are for telecommuters and typical subscribers.

In Canada, 3000 cell phones/day are being activated. One million new users will be wireless by the end of the year [Mercer 97]. Currently Canada has approximately three million users. The revenues are expected to be $750 million and the fraud estimate is $7.5 million, i.e. 1% of revenue. In the US, there are 25 million cellular phone users but fraudulent activity is estimated at $5 billion, representing a considerably larger percentage of revenue. This can be explained by the fact that wireless service has been available in the US longer than in Canada. Thus, criminals have had more time to develop sophisticated airwaves scanners that read the identification codes of the mobile phones as they are being transmitted electronically to the Base Station. The scanners work more than 1 kilometer away from the unsuspecting mobile phone user who is on the freeway or at a local mall. Once a number is picked up, it is programmed into a stolen mobile phone and rented out at $100/month. The unsuspecting user then gets billed for thousands of dollars in calls

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

they cannot identify. Inevitably, the mobile service provider absorbs the loss, passing the cost on to the general user base in terms of higher service tariffs.

This paper describes different types of fraud and approaches to combat them in section 2 and in section 3 we propose a multi-agent systems approach that would distribute the solution to fraud detection so that it may be more of a real-time solution that is effectively managed. Section 4 highlights some of the related issues and possible future work.

2. Types of Fraud and Deterrent Approaches

Mobile service providers face a difficult task in addressing fraud. A mobile network by its nature of being based on radio communications has many weak points that make it vulnerable to fraud -- radio signals are considerably easier to capture than signals on copper wire. One approach has been the move to digital networks rather than analogue networks. This makes the interception of digitally encoded MIN/ESN pairs through eavesdropping difficult but not impossible. In Europe, where the air interface standard is for a digital network, namely GSM (Global System for Mobility), fraud remains an issue. In North America, where the air interface is moving from analogue to digital, fraud is a critical problem costing nearly $6 billion a year. Three key types of fraud have been attacked by the mobile communications providers.

Tumbling Counter-Acted by Pre-Call Validation

Tumbling was the first kind of cellular fraud subscribers were faced with [Lo 95]. Since the first cellular switches in the US were not networked together, they could not exchange information. If a switch received a request from a phone outside of the call domain, it could not check the validity of the call. Thus, fraudsters programmed a random mobile id number from a different area into a phone and made up an ESN. After each call the ESN was increased by 1 number thus *tumbling* the number. Mobile communications companies counter-acted this approach by simply networking their switches and developing an IS-41 protocol by which the switches could exchange information. Thus, as a call arrived, a switch would first validate its MIN/ESN from its local database or from that of the subscriber's home switch. This approach has virtually eliminated tumbling as a type of fraud.

Cloning Counter-Acted by Profiling, Authentication, PINs & Smart Cards

Upon activation of a mobile phone, the MIN/ESN identification pair from a subscriber phone is transmitted to the nearest base station once every 15 minutes so that the switch knows the location of the phone and its nearest base station in the event of the user making a call. Around 1991, criminals began scanning the airwaves for MIN/ESN pairs with a simple device set at the right frequency near a busy traffic area such as an airport, highway, or a local shopping center [Lo 95]. The criminals then used the stolen MIN/ESN pair to reprogram another handset thus *cloning* the legitimate MIN/ESN pair and duping the switch. In this manner, they were able to make use of unlimited free service until the subscriber received their monthly bill.

Mobile communication companies started making use of a technique known as *profiling* to build a history of each user's calls through an analysis of network Call Detail Records (CDRs) and other measurements. Thus, in conjunction with the cellular company's billing system, there would be an application which continually profiles users. This profile could then be examined to determine if the use of a cellular phone is fraudulent or typical of the subscriber. The only drawback to this approach has been the inability of its use in real-time as CDRs are used to determine fraudulent activity and these are only generated at the end of a call. It may take a profiler 5 minutes to a day or more to detect that the use is fraudulent. In a *single* day, a cloned phone could be used for as much as two thousand dollars making it a direct loss for the phone company. Recently [Urqhart 96], par-

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

tial CDRs have been used in profiling in order to more quickly identify fraudulent calls that exceed a duration threshold.

A variety of techniques have been used to build user profiles including the tedious by hand construction of user profiles based on CDRs. Recently, neural networks have been used to classify profile data. Once a user profile pattern is formed by a neural network, anomalous and fraudulent patterns have been more easily detected due to their discrepancies with the expected pattern [Field and Hobson 97]. This approach, although effective, is not a real-time approach that is able to detect a fraudulent call as it happens.

Another approach to profiling is the use of data mining to detect indicators of fraud and then construct rules that are searched using thresholds [Fawcett and Provost 96]. Rules are used to construct profilers, weighted combinations of which are used to detect fraudulent activity. Call detail records are thus classified as legitimate or fraudulent based on the constructed rules. The account user profile is constructed using the individual profilers. On a daily basis, an account activity outline is constructed that consists of the subscriber calls and compared to the daily-threshold that is attempting to isolate unusual activity and high charges not typical of a particular subscriber. If the threshold is exceeded, an alarm is generated for the respective subscriber account and potential fraudulent activity noted. This method has been shown to work more successfully than checking the subscriber records by hand [Fawcett and Provost 96] and an accuracy of 94% is quoted for the approach.

The above approaches can also be combined with general fraud identification techniques. For example, multiple simultaneous use of a single number generate *collisions*. Furthermore, two calls occurring at unreasonably close times but are geographically dispersed based on a *velocity check* are detected easily. Collisions and velocity checks represent physical constraints that should not be violated and are often used by mobile service providers as fraud identification techniques.

Other approaches to combating cloning have included: Radio Frequency (RF) Fingerprinting, Authentication, Personal Identification Numbers (PINs), and Smart Cards. Every cellular phone could have a unique RF fingerprint [Lo 95]. The RF fingerprint is used in combination with the MIN/ESN to give the subscriber permission by the switch to place a call. This is a relatively new approach which is not supported by all the switches, but it is viewed as promising. Unfortunately, full implementation of this approach would require users to replace their existing handsets.

Authentication allows a switch to verify the identity of a cellular phone before processing a call. A secret key that is shared between the switch and the phone is transmitted over the airwaves. This key is used in combination with another authentication key registered to the subscriber as part of a calculation to ensure that the two keys match. This approach is known as private key encryption. Public key encryption requires the phone and the switch to exchange some parameters to decipher what the public key is. The public key is then used to encode the phone's identity and then transmit it to the switch. The public key approach avoids the storage of a private key in a database that may be accessed without authorisation. The public key is currently the preferred approach with mobile communication companies [Beller et al. 91 a; 91b; 91c]. One drawback to authentication technology is that the 28 million phones in North America do not have the authentication capability built-in. Thus, subscribers would have to replace their phones and the mobile communications companies would have to absorb the cost which could easily exceed $1 billion.

PINs are another parameter that mobile communication carriers make use of in combination with the MIN/ESN of a cellular phone. Phones are pro-

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

vided with security keys and a corresponding key at the exchange. In order to enable the authentication system the user has to enter her personal identification number to the SIM card on her handset. The exchange computer then sends out a random number whenever a call is made and the phone uses the security code to calculate a response [Lo 95], [Wong 95]. The drawback to this approach is that PINs can also be hacked and their transmission over the airwaves makes them as vulnerable to fraud as MIN/ESN identification parameters.

Smart cards are used in conjunction with the European GSM system to uniquely identify a user. These are cards with microchips built-in that provide unique subscriber identity information that are plugged into a cellular phone. Unfortunately, the cards themselves have also been susceptible to cloning [Lo 95].

Subscription Fraud Counter-Acted by Subscription Checking

Another type of fraud has criminals registering as legitimate subscribers, making extensive use of their phones especially by roaming and then simply disappearing and not paying their bills. [Field and Hobson 97], [Lo 95]. This type of fraud provides the criminal with the opportunity to make use of the network for a month before discovery. Assuming that subscription fraud activities have a pattern, companies are counter-acting this by screening subscribers and by maintaining records of fraudulent subscriptions. This information is shared between companies and criminals are prevented from signing up. No knowledge-based systems are currently used for screening subscribers but there is an opportunity to make use of intelligent profile builders here also.

Effective detection of fraudulent calls and other activity requires international cooperation as users roam from one country to another. Operators now fax other GSM partners information on high usage accounts and areas in order to spot fraud more quickly (historically it could take as long as three months before billing information would be sent to a partner).

3. A Multi-Agent Systems Approach

The current approaches to fraud detection are based on a multi-track approach in which the communications systems is protected using several different techniques. On one hand, the transmitted digital signal is encrypted in order to make it understandable only by its intended recipient(s). On the other hand, the fraud detection components are introduced in the control system or network in order to make real-time detection and/or disconnection of fraudulent calls possible. In one example, the common channel signalling (CCS) network's Service Switching Points (SSPs) are made capable of real-time detection of fraudulent calls while Service Control Point (SCP) databases are made capable of authenticating users by checking PIN, ESN, and Personal Communication System (PCS) numbers. In these systems the trade off between the amount of signalling and the required reaction speed may lead to centralized or distributed architectures for fraud detection elements. In any case, the fraud detection issue may be addressed in conjunction with the implementation of usage thresholds defining measures such as number of successful or failed calls, the duration of calls, restriction of available services, and the analysis of other call-related data [Ahimovic and Michaels 93].

It is clear to the authors from researching the problem of fraud in Personal Communication Systems that a multi-faceted approach is required. Mobility companies require a distributed approach to combat fraud since several points in the network can be attacked. First, a subscriber phone can be cloned. Second, a cloned phone cannot be detected until a reliable profile is constructed normally at a call processing center with access to switch information and billing records. Third, companies need to be able to detect a fraudulent call quickly and intercept it before significant charges are accumulated, i.e. calls in progress need to be monitored in real time for

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

fraudulent activity. We propose a multi-faceted distributed approach using embedded artificial intelligence techniques that combine Multi-Agent Systems and classification learning for fraud detection. We propose the use of Personal Communication Agents (PCAs), introduced in [Abu-Hakima et al. 96], in conjunction with rudimentary user profiles (UPs) of the subscriber's communication patterns. In addition, we propose the use of Mobility Network Agents (MNAs) to interact with PCAs. Finally, we propose the use of Fraud Breaking Agents (FBAs) that are similar in function to the profile constructing approaches that can be neural network-based or inductive algorithm-based [Fawcett and Provost 96], [Field and Hobson 97].

A typical subscriber is expected to make a call using a directory number (or a new number not in the directory that is logged as such). Such a subscriber will have various service categorisations for time-of-day and date. For example, a subscriber may call regarding business between 8:00 am and 6:00 pm and may call family and friends between 6:00 and 11:00 pm. They may then switch off their phone after 11:00 pm. A subscriber will also have a Personal Communication Agent acting as a personal assistant classifying and acting on their incoming messages as instructed or as it has learned by example. As a user makes a call, it is recorded by their User Profiler in a User Profile (UP). The Profiler is an intelligent process that is comparing the latest user communication with its historical information in the UP. If a fraudulent call is detected by the network, it is quickly communicated to the PCA. The PCA will make use of the UP to check the characteristics of this fraudulent call. If it is an atypical call for the subscriber, the PCA will make use of another means of communication (e.g. a pager, a desktop computer, etc.) to check with the subscriber. If the subscriber cannot be reached, the PCA will ask the network to disconnect the call otherwise it will await user instruction. The various elements of the multi-agent system are illustrated in Figure 1.

User Profiles

The UP would include a directory of numbers and addresses (such as email addresses) that a typical subscriber would keep. These directories would interwork with the user's seamless messaging applications (e.g. voice mail, email, fax, video mail) allowing them to direct any of their messages to particular users they would typically communicate with. Furthermore, typical parameters associated with a type of user communication would be monitored and logged. For example, the length of a user call and the assoicated number would be paired in their profile. Thus, if the user typically calls numbers in North America, a user profile would register this. Thus, as criminals clone a subscriber number and make calls to Asia, a UP can be examined for any instances to check if such calls are atypical of the subscriber.

Another aspect of the user profile would be the type of service categories that a subscriber fits. For example, a subscriber may only make use of their cellular phone in the evenings to call relatives and friends. Thus, they may have a service profile that switches off international calls between 7:00pm and 8:00am. If a subscriber phone is cloned and calls are made in the evening that are atypical of the user, the UP can be examined to detect fraud. A subscriber should also be permitted to examine their service categorisation and modify it if they wish. For example, a user may typically have used a cellular phone for evening calls to family and friends until they begin a business venture with others in an opposing time zone. The user should thus be permitted to expand the categories of service on demand as their needs dictate. Such customisation by the user has to be permitted through simple command-driven interfaces rather than cryptic programming.

Another possibility is to allow the user to cap the length of calls to logged phone numbers. A user may wish to typically limit the length of a cellular call to the average of 4 minutes given the cost of

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

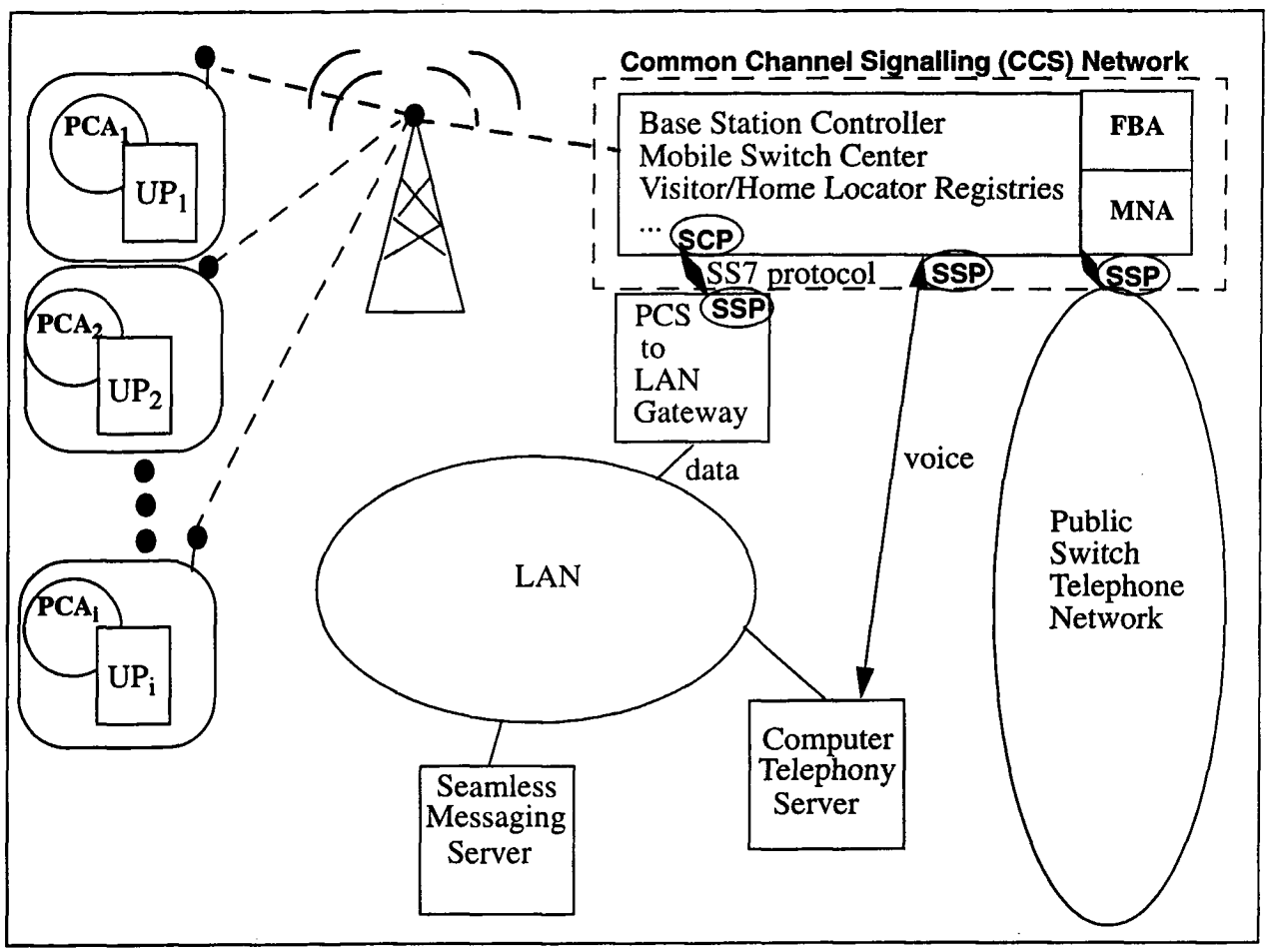

Figure 1. Multi-Agent System Configuration for Fraud in PCS.

air time. Thus, if a subscriber whose profile indicates that she typically only makes calls in North America for 4 minutes is suddenly making hour long calls, the system could suspect fraud and immediately attempt to restrict service through communication with the user through an alternate means (e.g. send them a pager or desktop email message that a call to the suspected number is going on and ask for permission to interrupt it).

Personal Communication Agents

Each subscriber is expected to have a PCA that holds a profile of the user's preferences in applications over heterogeneous networks (LANs, PCS, telephone, etc.) and with heterogeneous devices (pager, cellular phone, telephone, desktop, PDA, laptop, etc.). For example, a user may prefer to receive electronic mail on a cellular phone when they are telecommuting. For this to occur, the seamless messaging server makes use of a text to voice service agent to convert the message. The message is then delivered to the subscriber as a voice message. The original text message may have been filtered to a set of keywords first.

A personal communication agent is also expected to interact with a user's directories associated with their UP. It is possible for the PCA to log all the user's outgoing telephone call numbers, electronic mail addresses or fax numbers. As a result, the PCA has access to a very valuable tool in combating fraud, namely the numbers and addresses the user makes use of. From this information, a UP can be maintained which holds information such as who has been called, for how long, and when. The profiling of the information can be created in

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

a manner somewhat similar to the methods described in [Fawcett and Provost 96] or [Field and Hobson 97].

Mobility Network Agents

The mobility network requires intelligent processes that can interact with the subscriber PCAs. MNAs are expected to reside in the mobile switching center where the cellular calls are first processed in conjunction with home and visitor registeries used to allow the subscriber to roam. The role of the MNAs would be to interact with the PCAs to better profile a user. The MNA would provide the PCA with continuous access to the MNAs billing information for a particular subscriber. From this information the PCA can update its directory of calls made by the user (this allows it to check quickly if an "unusual call" is one that the user has called previously). As an MNA detects an "unusual call" it alerts the PCA. If the PCA concurs with the MNA about an "unusual call(s)", two actions can be taken: 1) alert the user on an alternate communication channel (desktop, land-line phone, pager, etc.); 2) monitor the call information until more evidence is collected that the call is indeed unusual (i.e. the pattern repeats itself, the bill is outrageous, etc.). The reason one may want two options is to allow the system to have a high probability of being correct, since repeated unwelcome warnings may backfire and force users to turn off their PCAs.

Another key aspect of the multi-agent system would be to include service profiling at the network end. As subscribers pick a service (emergency, leisure, business or business and leisure), a service profile is built through classification of what a typical profile of use fits that subscriber. Thus, if an emergency-only user starts calling overseas for hours at a time, the MNA should suspect fraud and warn the PCA.

Fraud Breaking Agents

"Unusual calls" can be detected by an FBA which also resides at the Mobile Switch Center. An FBA can make use of a single or combination of classification algorithms to detect the calling patterns typical of fraudulent cell phones: calls overseas that are hours long, phones that are on 24 hours/day, calls to known criminal centers or suspicious regions, etc. Based on FBA information, the MNA alerts the PCA to check the user profile and associated directories for any matching numbers and characteristics for the "unusual calls".

4. Issues and Future Work

Several issues can be raised in a multi-agent systems approach to fraud detection. Such an approach proposes the distribution of the problem so that some detection may be carried out faster at the subscriber end. However, an agents approach represents more of an open system versus the typical service provider's closed system. The open system introduces more points that may be vulnerable to fraud such as hacking User Profiles and falsely inserting numbers or clonning Personal Communication Agents. This can be deterred somewhat by using firewall type approaches to modifying the UP and the PCA thus ensuring that the user information is well guarded and accessible only to the user and the MNA. The inter-agent communication must also be encrypted to ensure its safety. Another issue is the distribution of subscriber records which may create privacy concerns for the user. A user may want their information to be wharehoused far from their local environment which again may be vulnerable. The complexity of inter-agent communication and interworking with call processing engines creates another layer of management on the network. This must be coupled with standardization efforts for multi-agent systems. Finally, the PCA/UP combination must not burden the user with requirements for hand-coded profiling. It is our experience that users dislike additional work in the use of their mobile phones.

Currently, we are in the process setting up the integrated testbed illustrated in Figure 1. We have implemented the first generation of seamless messaging agents as described in [Abu-Hakima et al. 96]. The second generation agents will support

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC

mobile agents that can gather information or cooperate with other geographically distributed agents. The implementation of the proposed fraud detection agents in conjunction with the User Profile will allow us to address the security aspect of network managment in seamless messaging.

References

[Abu-Hakima et al. 96] Abu-Hakima S., Liscano R., and Impey R., "Cooperative agents that adapt for seamless messaging in heteregeneous communication networks," IJCAI-96 Workshop on Intelligent Adaptive Agents, August 4, 1996, Portland, Oregon, Technical Report WS-96-04, AAAI Press, Menlo Park, California, pp.94-103.

[Ahimovic and Michaels 93] Ahimovic S.M. and Michaels J.M., "Services for tomorrow's PCS," Proceedings of the second IEEE international conference on Universal Personal Communications, October 12-15, 1993, Ottawa, Ontario, vol.1, pp.222-27.

[Beller et al. 91a] Beller M., Chang L.-F., and Yacobi Y., "Privacy and Authentication on a Portable Communication System," IEEE Global Telecommunications Conference, Globecom 91, Countdown to the New Millennium, Featuring a Mini-theme on: Personal Communication Systems (PCS), Phoenix, Arizona, December 2-5, 1991, vol.3, pp.1922-27.

[Beller et al. 91b] Beller M., Chang L.-F., and Yacobi Y., "Privacy and Authentication on a Portable Communication System," IEEE Journal on selected areas in communications, vol.11, no.6., August 1993, pp.821-829.

[Beller et al. 91c] Beller M., Chang L.-F., and Yacobi Y."Security for Personal Communication Services: Public-Key vs. Private Key Approaches," Proceedings of the second international symposium on personal, indoor, and mobile radio communications, PIMRC 92, 19-21 October, 1992, Boston, MA, USA, pp.26-31.

[Fawcett and Provost 96] Fawcett T. and Provost F., "Combining data mining and machine learning for effective user profiling," KDD-96: Proceedings of the second conference on knowledge discovery and data mining, August 2-4, 1996, Portland, Oregon, pp.8-13.

[Field and Hobson 97] Field S.D.H. and Hobson P.W., "Techniques for telecommunications fraud management," Accepted for publication in the proceedings of the third International Workshop on Applications of Neural Networks in Telecommunications, August, 1997.

[Handzel 94] Handzel M., "rules of the game: a set of guidelines can be used to model business processes across a wide variety of circumstances," Object Magazine, vol.4, no.1, March-April, 1994, pp.72-75.

[Lo 95] Lo J., "Fraud Busters: Who ya gonna call?," Telephony, vol.229, no.9, August 28, 1995, pp.22-26.

[McCulley and Rappaport 93] McCulley S.L. and Rappaport T.S., "Distributed real-time processing for cellular and paging traffic analysis, fraud detection and intelligent network wireless control," Proceeding of 1993 IEEE Vehicular Technology Conference, 18-20 May, 1993, Secaucus, NJ, USA, pp.891-6.

[Mercer 97] Mercer J., "Cellular phone fraud: when thieves get your number, scanners make cell-phone owners vulnerable to fraud," The Ottawa Citizen, January 22, 1997, pp.A1-2.

[Rice 96] Rice P., "SS7 Networks in a PCS World," Telephony, June 24, 1996, pp.138-146.

[Urquhart 96] Urquhart J.R., "Mining for gold," Telephony, June 24, 1996, pp.138-146.

[Wong 95] Wong K., "Fighting mobile phone fraud," Computer Fraud and Security Bulletin, January, 1995, pp.9-15.

EX-1014
AT&T SERVICES INC., CELLCO PARTNERSHIP D/B/A VERIZON WIRELESS,
AND NOKIA OF AMERICA CORPORATION v RIGHTQUESTION, LLC