

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

FORCEPOINT LLC,
Petitioner,

v.

WEBROOT, INC. AND OPEN TEXT, INC.,
Patent Owners.

IPR2023-00784
U.S. Patent No. 8,438,386

DECLARATION BY MICHAEL T. GOODRICH, PH.D.

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	SUMMARY OF OPINIONS.....	1
III.	BACKGROUND AND QUALIFICATIONS.....	2
	A. Education Background and Career History.....	2
	B. Publications and Patents.....	2
	C. Other Relevant Qualifications.....	4
IV.	UNDERSTANDING OF APPLICABLE LAW.....	7
	A. General Patent Law	7
	B. Obviousness.....	8
	C. Level of Ordinary Skill in the Art.....	10
V.	U.S. PATENT NO. 8,438,386.....	12
	A. Background Technologies for the ‘386 Patent.....	15
	B. Prior Art References.....	20
	1. Hubbard-495	20
	2. Hubbard-174	22
	3. Alperovitch.....	22
VI.	CHALLENGED CLAIMS.....	23
VII.	CLAIM CONSTRUCTION.....	24
VIII.	ANALYSIS OF HUBBARD-495 IN COMBINATION WITH HUBBARD-174 FOR CLAIMS 1-13, 15, 16, AND 18	25
	A. Independent Claim 1	25
	B. Claim 2	74
	C. Claim 3	76
	D. Claim 4	81
	E. Claim 5	82
	F. Claim 6	85
	G. Independent Claim 7	89
	H. Claim 8	100
	I. Claim 9	101

J.	Claim 10	104
K.	Independent Claim 11	107
L.	Claim 12	111
M.	Independent Claim 13	112
N.	Claim 15	123
O.	Claim 16	123
P.	Claim 18	126
IX.	ANALYSIS OF HUBBARD-495 IN COMBINATION WITH HUBBARD-174 AND ALPEROVITCH FOR CLAIMS 1-13, 15, 16, AND 18.....	127
A.	Independent Claims 1, 7, 11, and 13.....	127
B.	Dependent Claims 2, 3, 5, 6, 8, 9, 10, 12, 15, 16, and 18	130
C.	Dependent Claim 4.....	131
X.	SECONDARY CONSIDERATIONS	131
XI.	CONCLUSION	132

I, Michael T. Goodrich, Ph.D., declare under penalty of perjury as follows:

I. INTRODUCTION

1. My name is Michael T. Goodrich, Ph.D. I am a Professor in the Department of Computer Science at the University of California, Irvine. I have been retained on behalf of Forcepoint LLC (“Forcepoint”). I understand that Forcepoint is requesting that the Patent Trial and Appeal Board (“PTAB” or “Board”) institute an *inter partes* review (“IPR”) proceeding regarding claims 1-13, 15, 16, and 18 of U.S. Patent No. 8,438,386 (Ex.1001, “the ‘386 Patent”). I have been asked to analyze the ‘386 Patent, and to provide an assessment as to the validity of the ‘386 Patent in light of prior art.

2. For time spent in connection with this matter, I will be compensated in the amount of \$750 per hour. My compensation does not depend on the content of my opinions or the outcome of this case.

II. SUMMARY OF OPINIONS

3. Based on my education, experience, and review of the materials cited in this declaration, I have reached the following conclusions:

- Hubbard-495 in combination with Hubbard-174 renders obvious claims 1-13, 15, 16, and 18 of the ‘386 Patent
- Hubbard-495 in combination with Hubbard-174 and Alperovitch renders obvious claims 1-13, 15, 16, and 18 of the ‘386 Patent

III. BACKGROUND AND QUALIFICATIONS

4. Details of my education, work experience, and research are summarized in my curriculum vitae (“CV”) which is attached to this Report as Appendix A. I have summarized some of the relevant information which is on my CV herein.

A. Education Background and Career History

5. I received a Bachelor of Arts degree in Mathematics and Computer Science from Calvin University, in Grand Rapids, Michigan, in 1983, a Master of Science (M.S.) degree in Computer Science from Purdue University, in West Lafayette, Indiana, in 1985, and a Ph.D. in Computer Science from Purdue University in 1987.

6. I am a Distinguished Professor in the Department of Computer Science at the University of California, Irvine, where I have been a faculty member since 2001. The Distinguished Professor title is a campus-level distinction and is reserved for faculty who have achieved the highest levels of scholarship over the course of their careers. I was a professor in the Department of Computer Science at Johns Hopkins University from 1987-2001.

B. Publications and Patents

7. I have authored or coauthored over 350 publications, including several widely adopted books, such as Data Structures and Algorithms in Java, Data Structures and Algorithms in C++, Data Structures and Algorithms in Python, Introduction to Computer Security, and Algorithm Design and Applications. My

publications include contributions to data structures and algorithms, information visualization, networking, graph algorithms, computational geometry, distributed and parallel algorithms, cloud security, information security and privacy, and algorithm engineering. For example, I have published peer-reviewed research articles, including publications J-63, J-71, J-78, C-80, C-83, C-85, C-91, C-100, and C-108, on stopping viruses in email attachments and on authenticating data that is downloaded from untrusted repositories. I have also published peer-reviewed research articles on the use of parallelism in computer systems, including publications J-5, J-7, J-12, J-16, J-19, and J-24. Also, several of my textbooks include entire chapters on software engineering, and several of my publications, including Ch-8, Ch-12, J-36, J-45, C-53, C-62, C-63, C-64, C-208, are on algorithm engineering. Altogether, my publications have, according to Google Scholar, been cited over 16,000 times.¹ My research has been supported by grants from the Defense Advanced Research Projects Agency (“DARPA”), the National Security Agency (“NSA”), the Office of Naval Research (“ONR”), the Army Research Office (“ARO”), and the National Science Foundation (“NSF”).

8. I am a co-inventor on several U.S. patents, including U.S. Patent No. 7,257,711, “Efficient Authenticated Dictionaries with Skip Lists and Commutative

¹ See <https://scholar.google.com/citations?user=sSS9gSoAAAAJ>

Hashing,” which discloses secure distributed data authentication schemes. I am also co-inventor of U.S. Patent No. 7,299,219, “High Refresh-Rate Retrieval of Freshly Published Content using Distributed Crawling,” which discloses a technology for quickly retrieving website data that can change frequently, so as to be stored in a search engine. I am also co-inventor of U.S. No. Patent 8,681,145, “Attribute Transfer Between Computer Models Including Identifying Isomorphic Regions in Polygonal Meshes,” which teaches how to map one mesh-based computer model to another. In addition, I am co-inventor of U.S. Patent No. 9,152,716, “Techniques for Verifying Search Results Over a Distributed Collection,” which discloses a system for searching the Internet so as to produce verifiable search results that can be produced by a search engine.

C. Other Relevant Qualifications

9. In addition, I have consulting experience in matters involving algorithms, cryptography, machine learning, digital rights management, computer security, mobile devices (including power management), networking, software, video streaming, and storage technologies.

10. I am a Fellow of the American Association for the Advancement of Science (“AAAS”), a Fellow of the Institute of Electrical and Electronics Engineers (“IEEE”), and a Fellow of the Association for Computing Machinery (“ACM”), as well as being named as a Foreign Member of the Royal Danish Academy of Sciences

and Letters. I am also a recipient of a Fulbright Scholarship (for senior specialist service to University of Aarhus, Denmark). In addition, I am a recipient of the IEEE Computer Society Technical Achievement Award (“for outstanding contributions to the design of parallel and distributed algorithms for fundamental combinatorial and geometric problems”) and the Pond Award for Excellence in Undergraduate Teaching. Also, I am an ACM Distinguished Scientist.

11. My study of computer and network security topics began as an undergraduate student and continued in graduate school, where my Ph.D. research involved the study of computer system components operating in parallel. My study and interest in computer security continued after my Ph.D., as detailed above and in my CV. In addition, I have reviewed and evaluated research papers on computer security, including cryptography, beginning with work as an associate editor for Journal of Computer & System Sciences, as well as my service on program committees of peer-reviewed Computer Science conferences, including the Conference on Electronic Publishing and the Information Superhighway and the ACM Symposium on Theory of Computing (STOC), the latter for which I chaired and edited the conference proceedings.

12. I have taught courses at Johns Hopkins University, Brown University, and University of California, Irvine, at both the undergraduate and graduate levels. Topics of my courses have included computer security, algorithms, data structures,

information visualization, computer graphics, networking, algorithm engineering, computational geometry, and parallel processing. In addition, I have mentored twenty-five (25) PhD students over the years, who have written their PhD theses on topics in algorithms, data structures, information visualizations, networking, parallel processing, and computer security and privacy.

13. I have served as an editor on several technical journals, including Computational Geometry: Theory and Applications, Journal of Computer & System Sciences, Journal of Graph Algorithms and Applications, Int. Journal of Computational Geometry & Applications, and Information Processing Letters. I have also served on many program committees (PCs) for top conferences and workshops in Computer Science, including serving as PC chair in several instances. Examples include ACM Symposium on Computational Geometry (“SoCG”), ACM Symposium on Theory of Computing (“STOC”), Workshop/Symposium on Algorithms and Data Structures (“WADS”), Algorithm Engineering and Experimentation (“ALENEX”), which I cofounded with Dr. Catherine McGeoch in 1999), IEEE Symposium on Foundations of Computer Science (“FOCS”), ACM-SIAM Symposium on Discrete Algorithms (SODA), International Symposium on Graph Drawing (“GD”), International Colloquium on Automata, Languages, and Programming (“ICALP”), ACM Conference on Computer and Communications Security (“CCS”), European Symposium on Algorithms (“ESA”), IEEE

International Parallel and Distributed Processing Symposium (“IPDPS”), ACM Symposium on Parallel Algorithms and Architectures (“SPAA”), ACM Symposium on Advances in Geographic Information Systems (“GIS”), IEEE Symposium on Security and Privacy (“S&P”), IEEE International Conference on Big Data, IEEE International Conference on Data Engineering (“ICDE”), and International Symposium on Algorithms and Computation (“ISAAC”), and DAGS Conference on Electronic Publishing and the Information Superhighway.

IV. UNDERSTANDING OF APPLICABLE LAW

14. I am not an attorney. For the purposes of this declaration, I have been informed about certain aspects of the law that are relevant to my opinions. My understanding of the law was provided to me by the Petitioner’s attorneys.

A. General Patent Law

15. I understand that in order to determine the validity of a patent claim, the meaning and scope of that patent claim must be determined. I understand that in construing or interpreting a patent claim, the words of the claim are generally given their ordinary and customary meaning, which is the meaning that words would have had to a person of ordinary skill in the art (“POSITA”) at the time of the invention. The POSITA is deemed to read the claim term in the context of not only the particular claim in which the disputed term appears, but also in the context of the entire patent, including the other claims and the specification, as well as the

prosecution history of the patent. I further understand that an invalidity analysis requires that the claims (as construed) be compared to the prior art. If a claim was anticipated by or obvious over the prior art, the claim is invalid. I understand that the invalidity analysis is conducted based on the understanding of a POSITA at the time of the invention. I understand that prior art to the '386 Patent may include patents and printed publications that predate the alleged priority date of the '386 Patent.

B. Obviousness

16. It is my understanding that a patent claim is invalid for obviousness if the claimed invention would have been obvious at the time the invention was made to a POSITA. I further understand that a claimed invention may be invalid even if each and every limitation is not disclosed in a single prior art reference. I have been informed that, under the doctrine of obviousness, a claim may be invalid if the differences between the invention and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which the subject matter pertains. Obviousness, as I understand it, is based on the scope and content of the prior art, the differences between the prior art and the claim, the level of ordinary skill in the art, and secondary indicia of obviousness and non-obviousness to the extent they exist.

17. I understand that a claim can be found invalid as obvious if design incentives or market forces provided a reason to make an adaptation, and the invention resulted from the application of prior knowledge in a predictable manner. I understand that a claim can be found invalid as obvious if the claim would have been obvious because the substitution of one known element for another would have yielded predictable results to one of ordinary skill in the art at the time of the invention. I understand that a claim can be found invalid as obvious if the claim would have been obvious because the technique for improving a particular class of devices was part of the ordinary capabilities of a person of ordinary skill in the art (a “POSITA”). I understand that a claim can be found invalid as obvious if the claim would have been obvious because a particular known technique was recognized as part of the ordinary capabilities of one skilled in the art. I understand that a claim can be obvious in light of a single reference, without the need to combine references, if the elements of the claim that are not found in the reference can be supplied by the common sense of one of skill in the art.

18. Where an invention is a combination of known elements, I understand that I should consider whether a reason existed at the time of the invention that would have prompted one of ordinary skill in the art to combine the known elements in the way the claimed invention does. The reason could come from the prior art, the

background knowledge of one of ordinary skill in the art, the nature of the problem to be solved, market demand, or common sense.

19. I also understand that a finding that a POSITA would not have been able to practice the invention without undue experimentation or unpredictable results can defeat a claim of obviousness.

20. I have been informed and understand that in evaluating whether patent claims are invalid as obvious, objective indicia of non-obviousness can be considered. Such objective indicia can include commercial success, long-felt but unsolved needs, copying, praise, unexpected results, industry acceptance, failure of others, and skepticism by experts.

21. I also understand that it is impermissible to use the claimed invention itself as a blueprint for piecing together elements in the art. In other words, it is impermissible to use hindsight reconstruction to pick and choose among disclosures in the prior art to reconstruct the claimed invention.

C. Level of Ordinary Skill in the Art

22. I understand that there are multiple factors relevant to determining the level of ordinary skill in the pertinent art, including the educational level of active workers in the field at the time of the alleged invention, the sophistication of the technology, the type of problems encountered in the art, and the prior art solutions to those problems.

23. In determining the characteristics of a hypothetical POSITA of the ‘386 Patent at the time of the alleged invention, I considered several things, including the type of problems encountered in this field and the rapidity with which innovations were made. I also considered the sophistication of the technology involved, the educational background and experience of those actively working in the field, and the level of education that would be necessary to understand the ‘386 Patent. Finally, I placed myself back in the relevant periods of time and considered the state of the art and the level of skill of the engineers working in this field in those times.

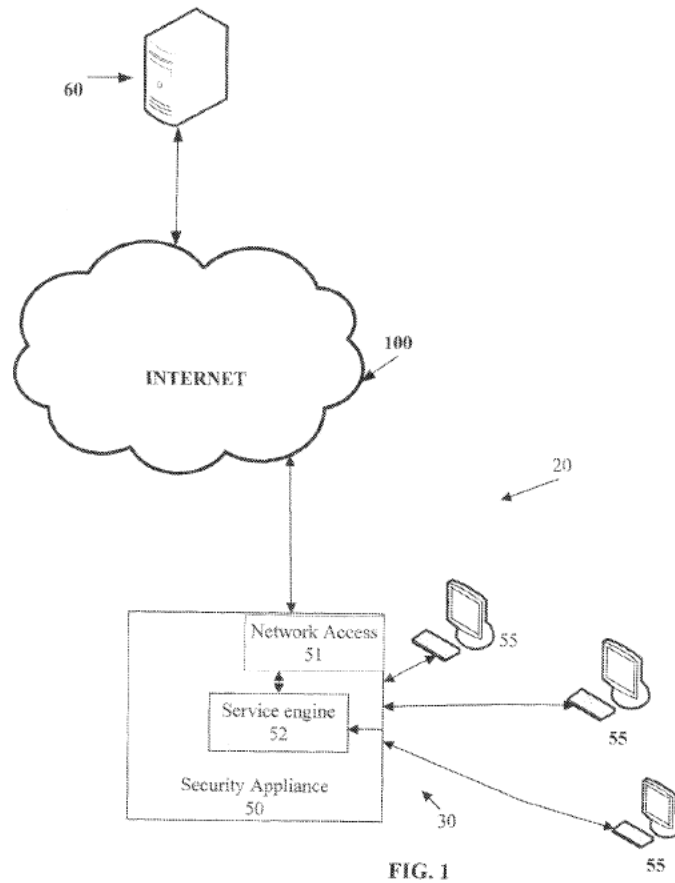
24. In my opinion, a POSITA for the ‘386 Patent would hold at least a four-year undergraduate degree in computer science or a closely related field and at least two years of experience in the field of computer and network security. Superior education would compensate for a deficiency in experience, and vice versa. For example, the ‘386 Patent identifies the field of its invention as relating “to assessing risk profiles of Internet resources.” Ex.1001, 1:22-24. More specifically, the ‘386 Patent states that its “invention is related to a system and method for developing a risk profile for an Internet resource by generating a reputation index, based on attributes of the resource collectively referred to as the reputation vector of the resource.” Ex.1001, 1:23-27. Also, the ‘386 Patent’s specification discloses the use of computer hardware and software (*see, e.g.*, Ex.1001, 3:49-63, 4:6-19).

25. I also note that my opinions provided in this Declaration would not change in view of any minor modifications to this level of skill.

V. U.S. PATENT NO. 8,438,386

26. U.S. Patent No. 8,438,386 was filed February 21, 2010, and issued May 7, 2013, and it cites provisional applications that were filed September 10, 2009 and April 21, 2009, respectively.

27. The '386 Patent relates to "[a] method and system for controlling access to an Internet resource." EX.1001, Abstract. As shown in Fig. 1, the system "comprises a local area network 30, the Internet 100 and an Internet service located at a remote server 60," where the local area network 30 "comprises a security appliance 50 and a plurality of client-side devices 55." *Id.*, 10:2-7.



Ex.1001, Fig. 1.

28. “Each of the client-side devices 55 preferably has a Web-browser for accessing the Internet from the client side device 55” and “[t]he security appliance 50 preferably comprises a network access 51 for permitting access to the Internet from the local area network 30, and a service engine 52 for determining if a requested Internet resource has a reputation index that meets a threshold established for the local area network 30.” Ex.1001, 10:10-16.

29. The '386 Patent generally discloses “transmitting a request for an Internet resource from an Internet-enabled client application from a client-side device of a local area network,” “receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet,” and “determining if a reputation index for the Internet resource is at or above a threshold value established for the local area network.” Ex.1001, 6:32-42; *see also id.*, 6:59-7:8. “The reputation index is calculated from a reputation vector which preferably includes a plurality of factors for the Web site comprising country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, to-level [*sic*] domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency.” *Id.*, 10:23-31. As seen in Fig. 3, “if the reputation index is not above the threshold,” then “access to the Web site is denied and a transmission of the denial is sent to the client-side device” whereas if “the reputation index for the Web site is above the threshold, then the access to the Web site by the client-side device is permitted by the security appliance.” *Id.*, 10:33-41.

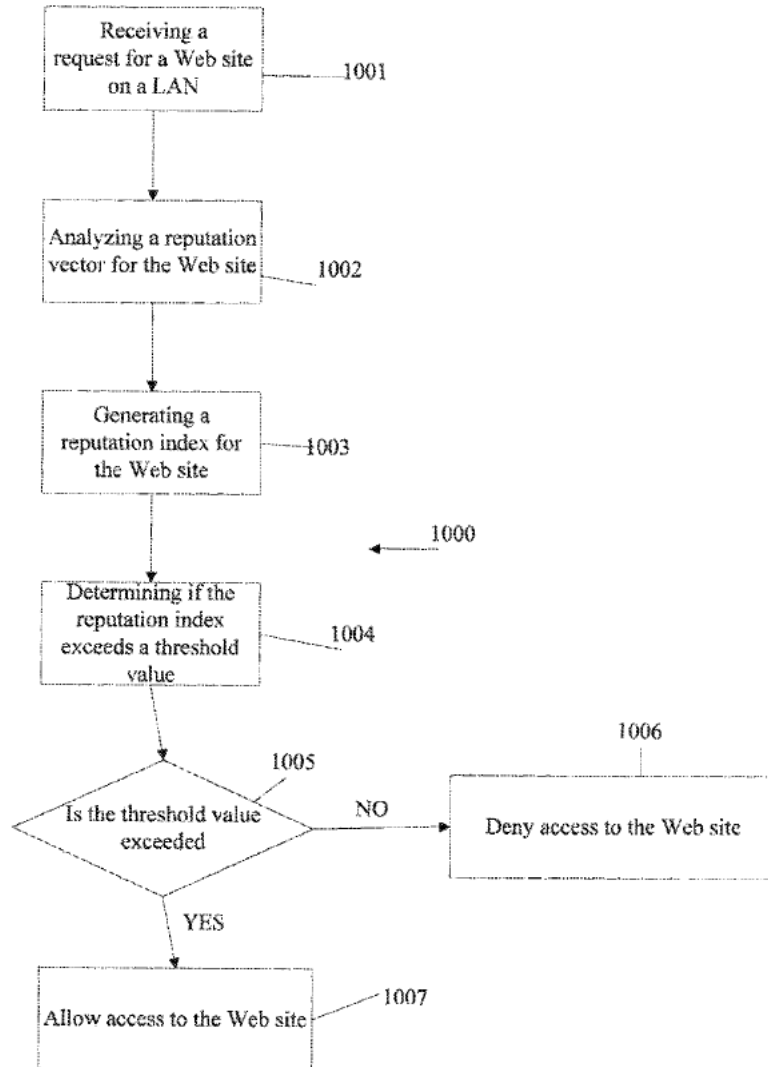


FIG. 3

‘386 Patent, Fig. 3.

A. Background Technologies for the ‘386 Patent

30. In this section, I review some background technologies for the ‘386 Patent, which a POSITA would understand.

31. As the '386 Patent defines it, and as a POSITA would understand, the “**Internet** is the worldwide, decentralized totality of server computers and data-transmission paths which can supply information to a connected and browser-equipped client computer, and can receive and forward information entered from the client computer.” Ex.1001, 3:44-48. As a POSITA would be aware, the Internet is comprised of hundreds of millions of computational devices connected in a world-wide network, which facilitates the sharing of information by sending and receiving data through this network.

32. The name, “Internet,” is derived from the fact that the Internet is a network of networks. That is, a POSITA would understand that computers connected in a network for an enterprise, school, or other organization, which is considered a “local” area, could be considered to be a “**local-area network**” (LAN), and a network that is comprised of multiple local-area networks could be called a “**wide-area network**” (WAN).

33. One of the primary tools for requesting and receiving information that is available on the Internet is to use a software application known as a **web browser**, which the '386 Patent defines as “a complex software program, resident in a client computer, that is capable of loading and displaying text and images and exhibiting behaviors as encoded in HTML (HyperText Markup Language) from the Internet, and also from the client computer's memory. Major browsers include MICROSOFT

INTERNET EXPLORER, NETSCAPE, APPLE SAFARI, MOZILLA FIREFOX, and OPERA.” Ex.1001, 3:66-4:5.

34. For example, a POSITA would know that documents and data can be stored, indexed, and retrieved in files that are encoded in the **HyperText Markup Language (HTML)**, which the '386 defines as “a method of mixing text and other content with layout and appearance commands in a text file, so that a browser can generate a displayed image from the file.” Ex.1001, 3:36-39. A POSITA would understand that HTML files can also include “**links**,” which are pointers to other files available on the Internet. Because of the web-like way in which such links can be visualized, the collection of data and links on the Internet is often referred to as the “**World-wide Web**.” For this reason, HTML files, or other files that are similarly encoded (e.g., in the XHTML or XML formats), so that they can be displayed in a web browser, are called “**web pages**.” In short, web browsers display web pages from the World-wide Web.

35. Links on the Internet are typically encoded using a format known as a **Uniform Resource Locator**, or **URL**, and the '386 Patent states, “URL or Uniform Resource Locator is a[n] address on the World Wide Web.” Ex.1001, 3:64-65. A POSITA would understand that a URL references a web resource. For example, the URL for my web page, which is encoded in HTML, and served from the University

of California, Irvine is <https://www.ics.uci.edu/~goodrich/index.html>. For instance, the '386 Patent defines a **web server** as follows:

Web-Server is a computer able to simultaneously manage many Internet information-exchange processes at the same time. Normally, server computers are more powerful than client computers, and are administratively and/or geographically centralized. An interactive-form information-collection process generally is controlled from a server computer, to which the sponsor of the process has access. Servers usually contain one or more processors (CPUs), memories, storage devices and network interface cards. Servers typically store the HTML documents and/or execute code that generates Web-pages that are sent to clients upon request. An interactive-form information-collection process generally is controlled from a server computer, to which the sponsor of the process has access.”

Ex.1001, 4:6-19.

36. Because networks, such as the Internet, can be so vast, they are often indexed using a tool known as a “**web crawler**,” which is a computer program that is used to search and index content and other information over the Internet. *See, e.g.*, https://web.archive.org/web/20090320030830/en.wikipedia.org/wiki/Web_crawler. These programs are commonly used to create entries for a **search engine** index, such as in the well-known GOOGLE service. The operation of a crawler typically involves reading HTML (or similar) documents to identify their content and links, collecting indexing information about them, and then following their links to repeat the process for other web pages. For example, the '386 Patent states, “[o]btaining the plurality of factors for the Internet resource preferably comprises accessing the Internet service, analyzing a plurality of HTML documents for the Internet resource,

and crawling a plurality of linked Internet resources of the plurality of HTML documents for Internet resource.” Ex.1001, 7:36-40.

37. A POSITA would know that computing devices on the Internet have addresses, known as **Internet Protocol (IP) addresses**, each of which is a sequence of bits that identifies that device. In IP version 4 (IPv4) these addresses are 32-bits long and in IP version 6 (IPv6) they are 128-bits long. These addresses are not suitable for humans to remember them, however, so to provide an easier way to refer to computing devices on the Internet, we often use **domain names**, which are strings such that each identifies a realm of autonomy, authority, or control, which is used to identify services or data, such as web pages or email. For example, the domain name for the web server for my web page at the University of California, Irvine is www.ics.uci.edu, and, as of this writing, its IP address, using a standard format for the 32 bits, is 128.195.1.88. Relatedly, a POSITA would understand that a **website** is a collection of web pages and related content identified by a common domain name and published on at least one web server. For example, the ‘386 Patent motivates its invention, e.g., by stating, “[a] Web user would want to be informed of the reputation of a Web site before visiting it, just as a lender would want to know the reputation, the financial reputation at least, of a borrower of the lender's money.” Ex.1001, 1:52-55.

38. A POSITA would understand that another design feature of the Internet is that it was designed to tolerate computer and communication failures. Indeed, the way data is sent on the Internet in discrete packets, rather than locked-down connections (as in the old telephone system), is because of this design choice. Further, a POSITA would be aware of several common techniques used on the Internet so that it can robustly tolerate computer and communication failures. For example, one common technique is to perform data **backups**, where data in a database or data store is sent (i.e., “backed up”) to a remote location, so that in the event that the computer for the database or data store is down or unavailable, its data can be recovered. See, e.g.,

<https://web.archive.org/web/20090304083309/en.wikipedia.org/wiki/Backup>.

Another common technique, which a POSITA would know, is **mirroring**, in which an Internet service is replicated (or “mirrored”) at a remote site, so that if the first site is inaccessible, the service can still be provided by the mirror site. See, e.g., https://web.archive.org/web/20081209010954/en.wikipedia.org/wiki/Mirror_site.

B. Prior Art References

1. Hubbard-495

39. U.S. Patent No. 9,654,495 (Ex.1004, “Hubbard-495”) was filed February 28, 2007, published June 5, 2008, and was issued May 16, 2017. I understand that Hubbard-495 is prior art to the ‘386 Patent.

40. Hubbard-495 discloses methods and system for “controlling access to web content.” Ex.1004, 2:33-61. A “reputation scoring module generates a reputation score or other reputation indicator of targeted content associated with the URL.” *Id.*, 27:38-40. “The reputation scoring module 331 may generate the scores based on information about the URL of the web page (e.g., scoring criteria).” *Id.*, 18:61-64. The disclosed factors include, for example, “[g]eo-location data associated with the URL (e.g., the URL host or host domain), the name server that provides the IP address of the URL host or its domain, the network or blocks of IP addresses associated with the host IP address or its domain, active content previously identified with reference to the host, the hosts’ domain, or the URL path, [and] historical data associated with the IP address of the URL host.” *Id.*, 18:61-19:23.

41. The reputation score may be “relied upon by the gateway server module 120 to determine under what circumstances, if any, the user will be allowed to access the content associated with each scored URL.” Ex.1004, 11:7-11. Hubbard-495 further discloses “receiving a request for at least one uniform resource locator (URL),” “generating an indicator of active content associated with the URL” that is “based on data associated with at least one component of the URL,” and “determining whether to allow the request based at least partly on the . . . indicator.” *Id.*, 2:34-40.

2. Hubbard-174

42. U.S. Patent No. 8,015,174 (Ex.1005, “Hubbard-174”) was filed February 28, 2007, published August 28, 2008, and was issued September 6, 2011. I understand that Hubbard-174 is prior art to the ‘386 Patent.

43. Hubbard-174 “relates to controlling access to the Internet by monitoring data received in response to a data request to the Internet.” Ex.1005, 1:7-9. A “gateway server module” is configured to “intercept the content sent from the first device to the second device and determine a category for the content based at least partially on a comparison of the intercepted content with data in a categorized URL/content database.” *Id.*, 2:1-5. A “policy module” contains “instructions for modifying content based on the category of the content,” and a “content filtering module” “modif[ies] at least some of the content based at least in part on the instructions.” *Id.*, 2:5-9. Hubbard-174 further describes that “modifying the content” consists of “removing at least some of the content based at least in part on a classification” and “sending the modified content to the requesting device.” *Id.*, 1:61-64.

3. Alperovitch

44. U.S. Patent Application Publication No. 2008/0175266 (Ex.1006, “Alperovitch”) was filed January 24, 2007 and was published July 24, 2008. I understand that Alperovitch is prior art to the ‘386 Patent.

45. Alperovitch discloses “[m]ethods and systems for assigning reputation to communications entities.” Ex.1006, Abstract. “Methods of assigning reputation to a communications entity can include: collecting data associated with entities originating communications,” “aggregating the collected data; analyzing the aggregated data to identify attributes respectively associated with entities originating communications; correlating the attributes to identify relationships between entities; updating a reputation associated with one or more entities based upon the relationship to one or more other entities identified by correlating the attributes; and, handling, communications based upon the updated reputation information.” *Id.*, [0009].

VI. CHALLENGED CLAIMS

46. I understand that Forcepoint challenges claims 1-13, 15, 16, and 18 of the ‘386 Patent. Claim 1 of the ‘386 Patent is reprinted below:

1. A method for controlling access to an Internet resource, the method comprising:

transmitting a request for an Internet resource from an Internet-enabled client application from a client-side device of a local area network, the Internet resource residing at a first server;

receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet;

determining if a reputation index for the Internet resource is at or above a threshold value established for the local area network, the reputation index generated from a reputation vector for the Internet resource, the reputation vector comprising a plurality of

factors for the Internet resource comprising security history, legitimacy, behavior, associations and location, wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block, wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency; and

transmitting a decision transmission to the Internet-enabled client application of the client-side device, the decision transmission allowing or denying access to the Internet resource.

VII. CLAIM CONSTRUCTION

47. For purposes of my analysis in this *inter partes* review proceeding, I have analyzed the terms appearing in the patent claims according to their “plain and ordinary meaning.” In determining the plain and ordinary meaning, I understand that I am to first give the terms their plain meaning that those words would have had to a POSITA in the context of the patent specification and claims. I understand that the structure of the claims, the specification, and the file history also may be used to better construe a claim insofar as the plain meaning of the claims cannot be understood. Moreover, treatises and dictionaries may be used, albeit under limited circumstances, to determine the meaning attributed by a POSITA to a claim term at the time of filing. I have followed this approach in my analysis, and for all of the

claim terms considered in this declaration, I have applied the plain and ordinary meaning of those terms.

VIII. ANALYSIS OF HUBBARD-495 IN COMBINATION WITH HUBBARD-174 FOR CLAIMS 1-13, 15, 16, AND 18

48. In my opinion, Hubbard-495, alone or in view of Hubbard-174, discloses or renders obvious claims 1-13, 15, 16, and 18 for the reasons set forth below.

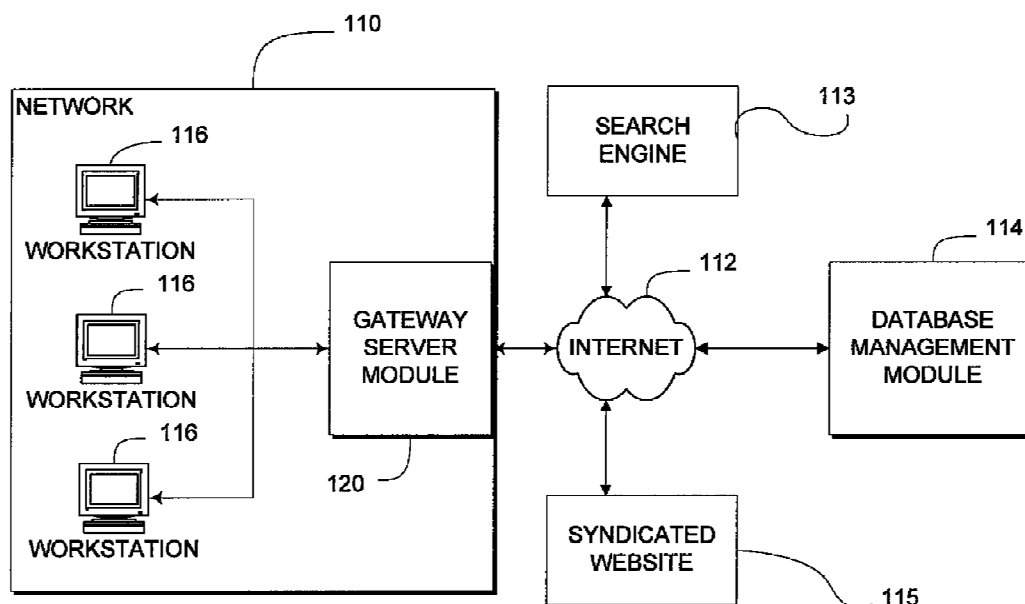
A. Independent Claim 1

49. Hubbard-495, alone or in view of Hubbard-174, discloses or renders obvious claim 1 for the reasons set forth below.

Motivation to Combine Hubbard-495 and Hubbard-174

50. A POSITA would have been motivated to combine Hubbard-495 and Hubbard-174 since both are analogous art to the '386 Patent at least because each is from the same field of endeavor, namely, methods and systems to determine conditions under which access to an Internet resource is allowed. Ex.1004, 2:33-61; Ex.1005, 1:7-9. For example, Hubbard-495 and Hubbard-174 have overlapping specifications and an inventor (Dan Hubbard) in common. *See, e.g.*, Ex.1004, Figs. 1-9, 4:38-67 5:8-22, 5:58-7:7, 7:46-8:14, 8:57-9:24, 11:41-14:51; Ex.1005, Figs. 1-9, 4:16-46, 4:58-65, 5:35-6:48, 7:57-8:25, 8:26-59, 8:60-12:3. A POSITA would understand that Hubbard-495 and Hubbard-174 disclose similar exemplary systems that could readily be combined. Ex.1004, 4:38-5:7; Ex.1005, 4:16-52. For example,

and as illustrated below, both disclose embodiments that include a gateway server module 120 connected through the Internet to a database management module. Ex.1004, 4:50-67; Ex.1005, 4:29-46. In those embodiments, workstations make requests for web content that may be allowed or blocked by the gateway server module. See Ex.1004, 6:34-36, 6:53-7:3; Ex.1005, 6:10-12, 6:27-45.



Ex.1004, Fig. 1; Ex.1005, Fig. 1.

51. In addition, a POSITA would have been motivated to combine the teachings of Hubbard-174 and Hubbard-495 as each reference teaches different functionalities of the same system. For example, Hubbard-495 discloses that “[t]he content filtering module 150 may be configured to parse the returned page and identify URL links in the returned search results and embedded items of content such as images or active content.” Ex.1004, 8:37-41. Hubbard-495 further discloses that

“the content filtering module 150 may also be configured to remove objectionable content from the returned search engine results.” *Id.*, 8:50-52. As Hubbard-495 does not elaborate further on how to remove objectionable content, it would have been natural for a POSITA to seek to combine Hubbard-495 with Hubbard-174, which discloses “removing certain portions of the data such as malicious or otherwise objectionable content.” Ex.1005, Abstract. Further, the two references share the same first-named inventor, as I mention above, and they were filed the same day, and listed the same assignee.

52. Similarly, Hubbard-174 describes a “scoring/classification system 186” that “provide[s] a score or classification . . . to web content” to “be delivered . . . to gateway server modules 120.” Ex.1005, 10:3-9. “[T]he gateway server module 120 may be configured to edit or identify returned URLs that are categorized as being ‘Malicious’ or ‘Spyware.’” *Id.*, 6:36-38. As Hubbard-174 does not elaborate further on how to score web content, it would have been natural to seek to combine Hubbard-174 with Hubbard-495 which discloses “generating a score indicative of the reputation, or likelihood that a web site associated with an uncategorized URL contains active or other targeted content.” Ex.1004, Abstract.

53. A POSITA would have had a reasonable expectation of success in combining Hubbard-495 and Hubbard-174, as doing so would have yielded nothing more than the predictable results of enhancing the common system of Hubbard-495

and Hubbard-174 by adding the functionalities described in each reference without the need for rearrangement or excessive modification of the system.

1[preamble]: “A method for controlling access to an Internet resource, the method comprising:”

54. In my opinion, Hubbard-495 discloses or renders obvious the preamble.

55. Hubbard-495 discloses a “method of controlling access to web content.” Ex.1004, 2:33-34. Hubbard-495 further discloses “receiving a request for at least one uniform resource locator (URL),” “generating an indicator of active content associated with the URL” that is “based on data associated with at least one component of the URL” and “determining whether to allow the request based at least partly on the at least indicator.” *Id.*, 2:34-40. Figure 25 from Hubbard-495 shows “one embodiment of a method 3000 of controlling access to web content based on the components of a requested URL.” *Id.*, 27:32-34. A POSITA would have understood that a URL (Uniform Resource Locator) references the web content or an “Internet resource” at that location. For example, the ‘386 Patent discloses the Internet resources of: “www.c.reditcan.cn” and “www.USmoney.GOV.” Ex.1001, 12:61-64. A POSITA would understand that the ‘386 Patent discloses web content as an Internet resource, stating, e.g., “Management of **internet** access, particularly to **Web sites**, has been accomplished in the past using ‘Content Filtering,’ where Web sites are organized into categories and requests for **Web content** are matched

against per-category policies and either allowed or blocked.”² *Id.*, 1:30-34. Also, the ‘386 Patent identifies a Web site as an exemplary Internet resource. *Id.*, 1:51-55, 4:36-38 (“The present invention focuses on the reputation of a Web site, or any Internet-based service or resource.”). Thus, I conclude Hubbard-495 discloses a method for controlling access to web content, such as disclosed also, e.g., in Fig. 25 excerpted below:

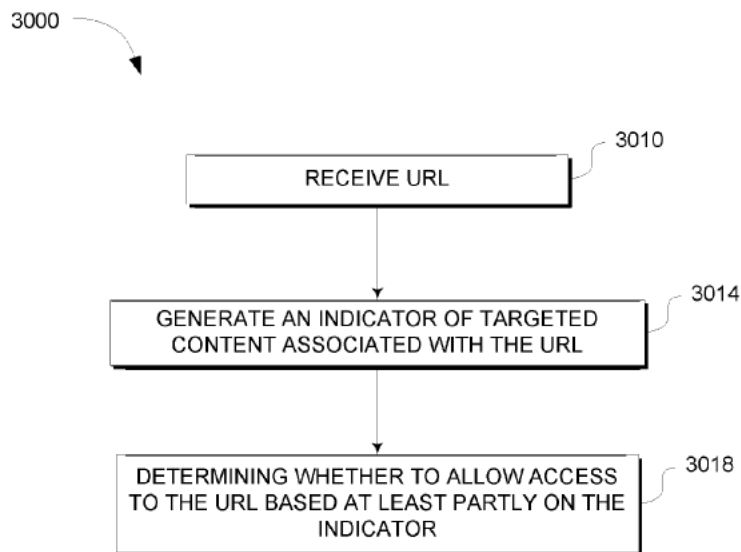


FIG. 25

² All emphasis added unless otherwise noted.

Hubbard-495, Fig. 25.

1[a]: “transmitting a request for an Internet resource from an Internet-enabled client application from a client-side device of a local area network, the Internet resource residing at a first server;”

56. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

57. As shown in FIG. 1, Hubbard-495 discloses an exemplary system that “includes a network 110. The network 110 may be a *local area network*, a wide area network, or some other type of network. The network 110 may include one or more workstations 116. The workstations 116 may be *various types of client computers that are attached to the network.*” Ex.1004, 4:39-44. “The client computers 116 may be desktop computers, notebook computers, handheld computers or the like.” *Id.*, 4:44-45. In my opinion, any one of the “workstations 116” or “client computers 116” that are attached to “the network 110”, which Hubbard-495 discloses may be a “local area network,” qualify as the claimed “client-side device of a local area network.”

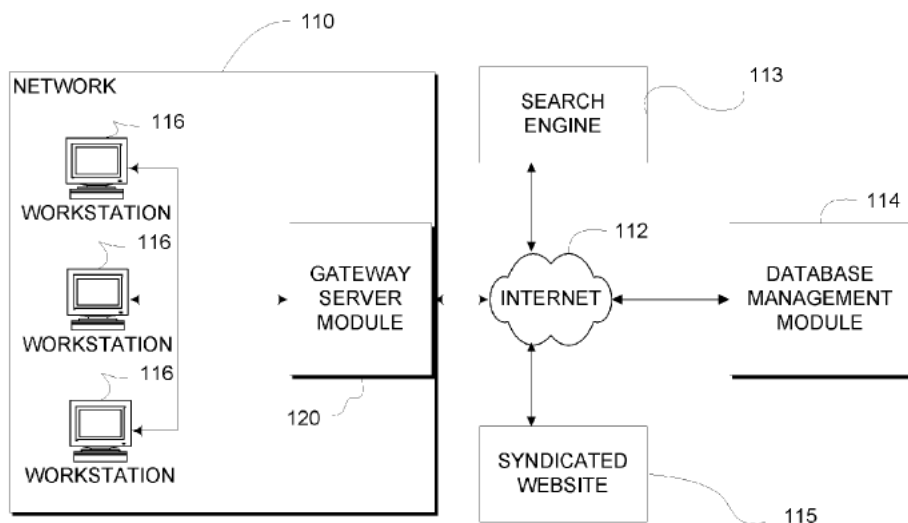


FIG. 1

Hubbard-495, Fig. 1.

58. As disclosed by Hubbard-495, the workstations or client computers are also equipped with Internet-enabled client applications. The client computers may “be loaded with operating systems that allow them to utilize the network through various software modules such as *web browsers, e-mail programs, or the like.*” Ex.1004, 4:46-49. A POSITA would have understood that web browsers and e-mail programs are inherently “Internet-enabled,” since they are used to access and display data sent over the Internet. In addition, Hubbard-495 teaches “the workstation 116 requests a URL from the Internet 112.” Ex.1004, 22:53-54. For example, the workstation 116 may include “one or more URL request modules 136.” *Id.*, 5:16-

17, 5:35-37. The functions of “the URL request modules 136 are performed by software applications such as web browsers, with Internet Explorer®, Mozilla Firefox, Opera, Safari, being examples of browsing software well-known in the art” but “[t]hese functions may also be performed by other types of software such as web feed readers, instant messaging programs,” “or some other applications.” *Id.*, 5:38-67.

59. Hubbard-495 further discloses that the workstations 116 “may be in communication with one or more search engines 113 via the Internet 112” and that “the search engine 113 is typically a web site which receives user requests for information and websites which are relevant to the parameters of the submitted user requests” such as “Google, Yahoo, and Ask.com.” Ex.1004., 5:1-7. A POSITA would have understood that the workstations 116 or client computers 116 would use web browsers or similar applications to access the search engines on the Internet. *See also, e.g., id.*, 4:7-10 (“As used herein, potentially executable web content generally refers to any type of content that includes instructions that are executed by a web browser or web client 10 computer.”), and 5:8-10 (“In the example system of FIG. 1, one or more of the workstations 116 may be in communication with one or more syndicated websites 115 via the Internet 112.”) In my opinion, the “various software modules such as web browsers, e-mail programs, or the like” installed on

the workstations 116 would qualify as the claimed “Internet-enabled client application from a client-side device of a local area network.”

60. Hubbard-495 discloses a “network request” that includes “user requests for content via the search engine 113 or requests sent via a web feed reader to the syndicated website 115.” Ex.1004, 5:51-54. The URL request module 136 may “receive a data request from either a user or some system process, and send the request via the gateway server module 120 to retrieve the file and/or content associated with the request.” *Id.*, 5:40-43. Hubbard-495 further discloses that the request “may be a search engine request made to the search engine 113.” *Id.*, 5:43-45. The search engine 113 may “return one website or a list of websites (and URL links leading to those websites)” in response to the search. *Id.*, 5:45-47. The request may also be “a content request made to the syndicated website 115” which may return “one URL link or a list of URL links to the requested content.” *Id.*, 5:48-51. A POSITA would have understood that “a search engine request” and a “content request made to the syndicated website 115” would be a request for an Internet resource.

61. These requests are made by the workstation 116, which “may include a workstation module 130.” Ex.1004, 5:16-17. The workstation module 130 may also include “one or more URL request modules 136.” *Id.*, 5:35-37. Further, as seen in FIG. 1, the search engine 113 and syndicated website 115 is accessible via the

Internet 112. In my opinion, the request sent by the URL request module 136 of the workstation module 130 of the workstation 116 would qualify as the claimed “transmitting a request for an Internet resource from an Internet-enabled client application from a client-side device of a local area network.”

62. As discussed, the workstations 116 may be in communication with “one or more search engines 113 via the Internet 112” with the search engine 113 “typically a web site which receives user requests for information and websites which are relevant to the parameters of the submitted user requests” such as “Google, Yahoo, and Ask.com.” Ex.1004, 5:1-7. The search engine 113 may “return one website or a list of websites (and URL links leading to those websites)” in response to the search. *Id.* at 5:45-47. Alternatively, the request may also be “a content request made to the syndicated website 115” which may return “one URL link or a list of URL links to the requested content.” *Id.*, 5:48-51. A POSITA would have understood that a “website or a list of websites (and URL links leading to those websites)” and “URL link or a list of URL links to the requested content” would be examples of “a request for an Internet resource,” since they each comprise a request for a resource on the Internet. It was well known that webpages would reside at a web server, for example. A POSITA would have understood that webpages on a web server would be an “Internet resource residing at a first server.” This understanding is likewise confirmed in Hubbard-495, e.g., *id.*, 5:8-11 (“In the

example system of FIG. 1, one or more of the workstations 116 may be in communication with one or more syndicated websites 115 via the Internet 112. The syndicated website 115 offers a web feed of content from the website.”). In my opinion, the search result content on a search engine 113 and/or website content on a syndicated website 115 (being requested by the workstation 116) would qualify as the “Internet resource residing at a first server.”

1[b]: “receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet;”

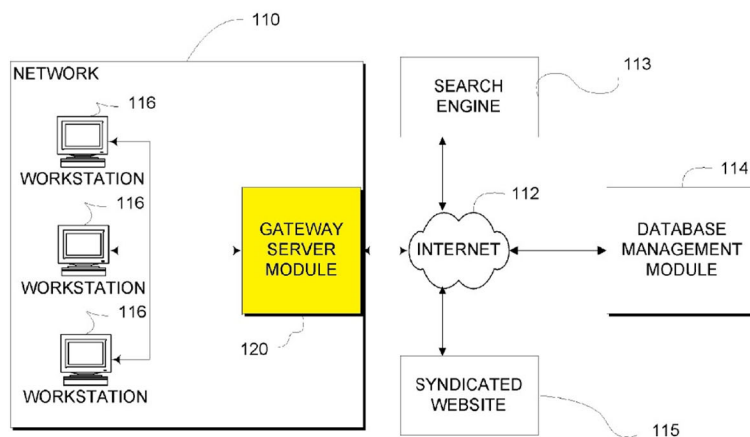
63. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

64. The ‘386 Patent describes the security appliance as “control[ling] access to the Internet by each of the plurality of client-side devices.” Ex.1001, 5:59-61. The ‘386 Patent security appliance “preferably comprises a network access 51 for *permitting access to the Internet* from the local area network 30, and a service engine 52 for determining if a requested Internet resource has a reputation index that meets a threshold established for the local area network 30.” *Id.*, 10:12-16.

65. Hubbard-495 discloses a gateway server module 120 that “may be used to analyze incoming and outgoing web traffic and to make various determinations about the impact the traffic may have on the workstations 116.” Ex.1004, 6:14-17. Hubbard-495’s gateway server module 120 is also security-oriented, as the gateway

server module may also “be integrated with some firewall hardware or software that protects the network 110 from unauthorized intrusions from the Internet 112.” *Id.*, 6:26-30. The gateway server module 120 may also include a policy module 142 that may be “used to **implement network policies** regarding how certain content will be handled by the gateway server module 120 or by a firewall or some other security software installed within the network 110.” *Id.*, 6:53-58.

66. Hubbard-495 discloses that the gateway server module 120 “may reside at the edge of the network 110 so that traffic sent to and from the Internet 112 may pass through it on its way into or out of the network 110.” Ex.1004, 4:51-54.



Ex.1004, Fig.1 (annotated).

67. As seen in FIG. 1, the gateway server module 120 is part of the network 110, which as discussed above with respect to element 1[a] may be a local area network. Further, Hubbard-495 discloses, that “the gateway server module 120 may

be a standalone hardware appliance.” Ex.1004, 6:30-33. In my opinion, Hubbard-495’s gateway server module 120 qualifies as the claimed “security appliance of the local area network.”

68. Hubbard-495 discloses that the URL request module 136 of the workstation 116 “may receive a data request from either a user or some system process, and send the request *via the gateway server module 120* to retrieve the file and/or content associated with the request.” Ex.1004, 5:40-43. The gateway sever module 120 “may *receive* file uploads and downloads and *URL requests* from the workstation module 130.” *Id.*, 6:20-22. As discussed above with respect to element 1[a], Hubbard-495 discloses that the request sent by the workstation 116, such as the URL request, may be a request for an Internet resource. In my opinion, Hubbard-495 discloses or renders obvious the claimed “receiving the request for the Internet resource at a security appliance of the local area network.”

69. Hubbard-495 further discloses that the gateway server module 120 is in “two way communication with the workstation 116” and “also in two-way communication with the Internet 112.” Ex.1004, 6:19-23. “Thus, *requests originating within the workstations 116 of the network 110 may be required to pass through the gateway server module 120 as they proceed to the Internet.*” *Id.*, 6:24-26. A POSITA would have understood that requiring requests to pass through the gateway server module as they proceed to the Internet would necessarily mean that

the gateway server receives the request *before* it continues on to the Internet. In my opinion, Hubbard-495 thus discloses or renders obvious the claimed “*receiving the request* for the Internet resource at a security appliance of the local area network *prior to transmission of the request over the Internet.*”

1[c]: “determining if a reputation index for the Internet resource is at or above a threshold value established for the local area network, the reputation index generated from a reputation vector for the Internet resource, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location, wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block, wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency; and”

70. In my opinion, Hubbard-495 discloses or renders obvious this limitation for the reasons set forth below.

1[c1]: “determining if a reputation index for the Internet resource is at or above a threshold value established for the local area network,”

71. The ‘386 Patent describes the reputation index as “a score which can be used with empirically developed threshold values to block access where the reputation index is deemed to be too low to be safe.” Ex.1001, 9:37-41.

72. Hubbard-495 discloses that a “URL analysis module 149” may “examine requested URLs” and “generate a *score* or other data indicative of whether the URL is likely to reference targeted (e.g., active) content.” Ex.1004, 9:25-29. Hubbard-495 further discloses that the URL analysis module 149 “may comprise a reputation scoring module (such as an embodiment of the reputation scoring module 331 of FIG. 12) that generates *a score* or other data representative of a reputation of the URL.” *Id.*, 9:29-33. The policy module 142 “may be configured to *block access to URLs with low scores (e.g., indicative of poor reputation or higher likelihood of active or other targeted content)* where the applicable policy might otherwise allow access to the URL.” *Id.*, 11:36-40. Thus, Hubbard-495 discloses a “score” representative of a reputation of the URL that can be used to block access to URLs with low scores. In my opinion, the reputation score disclosed by Hubbard-495 would qualify as the claimed “reputation index.”

73. Hubbard-495 further describes that “the gateway server module 120 determines whether to allow access to the URL *based at least partly on the reputation score.*” Ex.1004, 27:51-53. Additionally, “the reputation scoring module 331 may provide the reputation score to the content scoring module 330, which may determine a category of the requested URL at least partly based on the reputation score.” *Id.*, 27:54-57. “The gateway server module 120 may rely upon the determined category in combination with the score associated with the URL

when determining whether to allow the user access to the website associated with the requested URL.” *Id.*, 27:58-61. “Each reputation score may then be unique to each website URL and relied upon by the gateway server module 120 to determine under what circumstances, if any, the user will be allowed to access the content associated with each scored website URL gathered by such types of web feed reader applications.” *Id.*, 10:64-11:2. As shown, Hubbard-495 contemplates determining whether access is allowed to the requested URL based on the reputation score.

74. Hubbard-495 also discloses that the policy module 142 inside the gateway server module 120 “can be configured to implement different policies depending on the score, the user, or any other applicable parameters.” Ex.1004, 11:27-29. The policy module 142 may be “used to *implement network policies* regarding how certain content will be handled by the gateway server module 120 or by a firewall or some other security software installed within the network 110.” *Id.*, 6:54-58. The policy module 142 “may be configured to *block access to URLs with low scores (e.g., indicative of poor reputation or higher likelihood of active or other targeted content)* where the applicable policy might otherwise allow access to the URL.” *Id.*, 11:36-40. It would have been obvious to a POSITA that the gateway server module 120 and the policy module 142 of Hubbard-495, being part of the same local area network as the client workstations 116, would have a “network policy” that would apply within the network 110, which, as discussed above with

respect to element 1[a], may be a local area network. *Id.*, 4:38-41. For example, as shown in Fig. 3, the policy module 142 of the gateway server module 120 is shown as being a part of the local-area network of the client workstation 116; hence, a POSITA would find it obvious that the policy module 142 could be used to establish policies for the local area network. *See, e.g.*, Fig. 3, which I excerpt below:

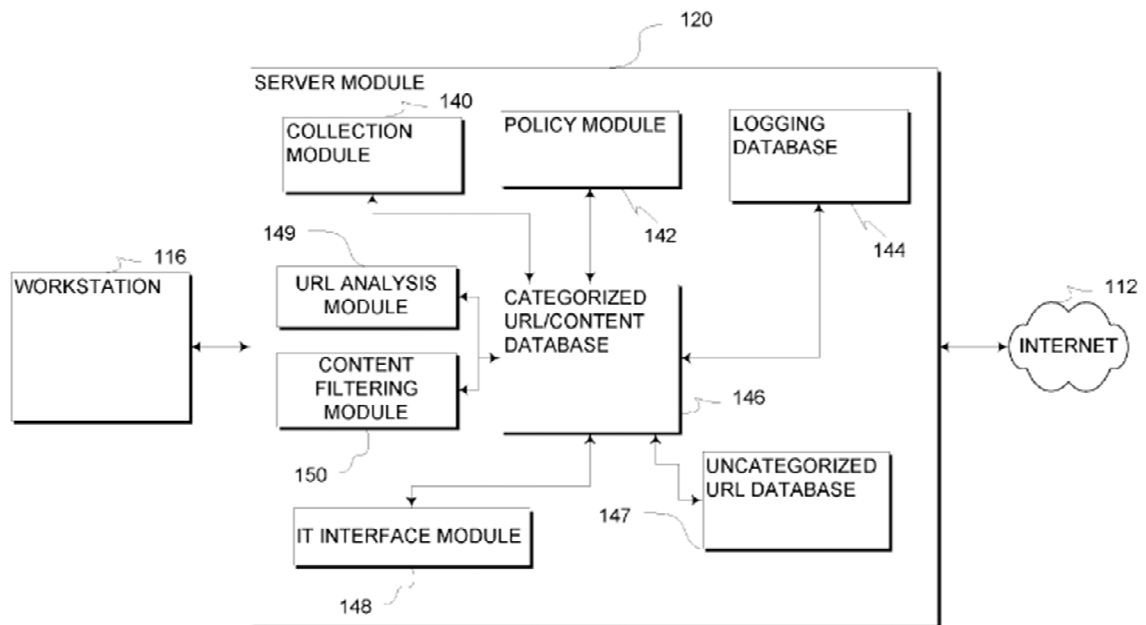


FIG. 3

Ex.1004, Fig. 3.

75. Further, Hubbard-495 describes that the URL analysis module 149, which is also a part of the gateway server module 120, may generate a “*sufficiently high*” reputation score (e.g., indicative of a good reputation or low likelihood of active or other targeted content) for an uncategorized URL so that the policy module 142

allows access to the uncategorized URL when the policy module 142 would otherwise not allow access to the URL.” Ex.1004, 11:27-36. A POSITA would have understood that determining whether a reputation score is “sufficiently high” for the particular local area network on which the gateway server module is located would necessarily involve comparing it to a threshold value.

76. Further, the policy module 142 “may be configured to provide the system guidance on how to handle URL requests for categorized URLs,” such as “disallow[ing] URL requests that are categorized as being ‘Malicious.’” Ex.1004, 6:58-63. Hubbard-495 contemplates the use of thresholds in the scoring process, where, e.g., “categories corresponding to each score above a given *threshold* are identified with the URL.” *Id.*, 21:1-4. *See also, e.g., id.*, 19:45-51, 19:55-61, 20:65-21:1 (“In one embodiment, *URLs with scores above a threshold* are identified with a category, e.g., malicious, regardless of, or in addition to, the category identified by content analysis.”), 21:1-10.

77. In my opinion, therefore, Hubbard-495 discloses or renders obvious “a threshold value established for the local area network.”

1[c2]: “the reputation index generated from a reputation vector for the Internet resource, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location,”

78. In the ‘386 Patent, the reputation index is generated “based on attributes of the resource collectively referred to as the reputation vector of the resource.”

Ex.1001, 1:24-28. The ‘386 Patent also describes the “collection of reputation-relevant features referred to as the reputation vector.” *Id.*, 4:61-62. “The reputation vector includes a plurality of factors for the Web site comprising at least one or more of country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency.” *Id.*, 5:23-30. Thus, a POSITA would understand that a reputation index in the ‘386 Patent is a value or score generated from a reputation vector, which comprises a plurality of factors.

79. Hubbard-495 discloses that reputation scores may be generated “based on information about the URL of the web page (e.g., scoring criteria)” such as:

one or more of URL content strings, Geo-location data associated with the URL (e.g., the URL host or host domain), the name server that provides the IP address of the URL host or its domain, the network or blocks of IP addresses associated with the host IP address or its domain, active content previously identified with reference to the host, the hosts' domain, or the URL path, historical data associated with the IP address of the URL host, creation date or expiration date associated with the registration of the URL host, whether search engine results exist for the URL or the URL host, certificate details associated with the URL (e.g., for secure (such as HTTPS) access schemes), meta tag categories associated with other URLs associated with the URL host (e.g., such as in the URL database 180 or the processed web page properties database 324), categories associated with other URLs of

the URL host (e.g., in the URL database 180), a fingerprint (e.g., a digest or hash) associated with the content of the URL for comparison to such data of previously categorized URLs (and to generate reputation data depending on whether URLs with the same fingerprint had targeted content), and scoring based on site (host or domain) access statistics (e.g., generate higher reputation scores for frequently visited sites than for infrequently visited sites because frequently visited sites (for other URLs that do not have targeted content) are less likely to have targeted content in newly visited URLs).

Ex.1004, 18:53-19:22.

80. A POSITA would have understood that these listed examples would be relevant to the reputation of the URL. As discussed above with respect to element 1[a], a POSITA would understand that a URL constitutes an Internet resource. In my opinion, this information comprising a plurality of factors about the URL (e.g., scoring criteria) on which a reputation score may be based would qualify as the claimed “reputation vector comprising a plurality of factors for the Internet resource.”

81. Hubbard-495 discloses that “the reputation scoring module 331 may generate the scores based on information about the URL of the web page (e.g., scoring criteria)” as listed above. Ex.1004, 18:53-19:22. In my opinion, this would qualify as the claimed “the reputation index generated from a reputation vector for the Internet resource.”

Security History

82. Hubbard-495 discloses that “information about the URL of the web page (e.g., scoring criteria)” may include “historical data associated with the IP address of the URL host.” Ex.1004, 18:53-19:22. It would have been obvious to a POSITA that “historical data associated with the IP address of the URL host” could include historical data about any of the “information about the URL of the web page (e.g., scoring criteria)” and that a number of Hubbard-495’s examples include security data, such as “certificate details associated with the URL (e.g., for secure (such as HTTPS) access schemes),” “active content previously identified with reference to the host, the hosts' domain, or the URL path,” and “a fingerprint (e.g., a digest or hash) associated with the content of the URL for comparison to such data of previously categorized URLs (and to generate reputation data depending on whether URLs with the same fingerprint had targeted content).” *Id.* For example, as of this writing, the certificate associated with the uci.edu domain for use with the HTTPS protocol has the following historical information (and a POSITA would

understand similar security history information would have been included in digital certificates for the HTTPS protocol as of the priority date of the ‘386 Patent):

Certificate Viewer: scpa-p-web1.ucomm.uci.edu

General

Details

Issued To

Common Name (CN)

Organization (O)

Organizational Unit (OU)

scpa-p-web1.ucomm.uci.edu

University of California, Irvine

<Not Part Of Certificate>

Issued By

Common Name (CN)

Organization (O)

Organizational Unit (OU)

InCommon RSA Server CA

Internet2

InCommon

Validity Period

Issued On

Expires On

Monday, March 6, 2023 at 4:00:00 PM

Wednesday, March 6, 2024 at 3:59:59 PM

Fingerprints

SHA-256 Fingerprint

SHA-1 Fingerprint

6E 2A 84 60 A9 0B 03 EA 8D 2F 2C 22 22 A2 3C 52 B7 11 BF 83 4C DE FF D8 35 3D C0 8B 11 F0 CF B5 01 7E E2 BA A4 2A 92 08 A6 9F B1 76 87 5C 58 6E E4 58 B8 D7

A POSITA would have known that “active content previously identified with reference to the host, the hosts' domain, or the URL path” would be an item that flags security risks, and that in order to know that the content was “previously identified with reference to the host, the hosts’ domain, or the URL path” would require historical information about the active content. In my opinion, this would qualify as

the claimed “security history” factor of the “plurality of factors for the Internet resource.”

Associations

83. Hubbard-495 discloses that “information about the URL of the web page (e.g., scoring criteria)” may include “meta tag categories *associated with other URLs associated with the URL host* (e.g., such as in the URL database 180 or the processed web page properties database 324), categories *associated with other URLs of the URL host* (e.g., in the URL database 180),” and “a fingerprint (e.g., a digest or hash) associated with the content of the URL for comparison to such data of *previously categorized URLs* (and to generate reputation data *depending on whether URLs with the same fingerprint had targeted content*).” Ex.1004, 18:53-19:22. Hubbard-495 also discloses “the reputation scoring module 331 may assign scores to entire subnetworks” which may be “based on the relative number of URLs in particular networks or domains that have a particular category. For example, if 95% of the URLs for a particular network in the URL database 180 are classified as having malicious active content, *new URLs may be given a score indicative of the likelihood that other URLs referencing servers in the same network are also likely to have such targeted content.*” *Id.*, 20:6-21. The ‘386 Patent states that associations may include, for example, “the number of sites linking into the site, the reputations of the linked in sites and the reputations of the linked-to sites.” Ex.1001,

9:19-21. A POSITA would have understood that “associations” would be information about *other* Internet resources that are not the requested Internet resource, such as the URLs, URL hosts, or networks disclosed above, that nevertheless have some relationship or component in common with the requested Internet resource, such as the meta tag categories, categories, fingerprints associated with content, and targeted or active content disclosed above. Hubbard-495 thus discloses that “scoring criteria” may include “associations” for the Internet resource. In my opinion, this information disclosed above would qualify as the claimed “associations” factor of the “plurality of factors for the Internet resource.”

84. As the elements of “location,” “legitimacy,” and “behavior” have further limitations in the claim, I address them below, and incorporate that analysis by reference here. In my opinion, Hubbard-495 discloses or renders obvious “the reputation index generated from a reputation vector for the Internet resource, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location.”

1[c3]: “wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block,”

85. First, Hubbard-495 discloses or renders obvious determining a reputation score based on country of domain registration. Hubbard-495 discloses that “information about the URL of the web page (e.g., scoring criteria)” may include

“***Geo-location data associated with the URL*** (e.g., the URL host or host domain),” and “the network or blocks of IP addresses associated with the host IP address or its domain.” Ex.1004, 18:53-19:22. A POSITA would have understood that “the network or blocks of IP addresses associated with the host IP address or its domain” could disclose the country of an internet protocol address block. Hubbard-495 further discloses that “entire subnetworks (e.g., all addresses matching 128.2.*.* may have a particular score)” and that such “networks or subnetworks help identify a web site as ***being based in a particular country*** or with a particular service provider.” *Id.*, 20:6-11. In addition, Hubbard-495 discloses that reputation may be based on “information about the registrar of the site (e.g., the domain name of a URL host)” which may include “the identity of the registrar, the ***registrant*** (including ***address and contact information of the registrant***).” *Id.*, 20:22-32. A POSITA would have known that the registrant address and contact information could provide location information about the URL. In my opinion, the above disclosures would qualify as the claimed “location factor of the plurality of factors for the reputation vector.”

86. Hubbard-495 discloses that reputation may be based on “information about the registrar of the site (e.g., the domain name of a URL host)” which may include “the identity of the registrar, the registrant (including address and contact information of the registrant), . . . and ***any other registrar data including data***

provided by the WHOIS protocol (e.g., Internet Engineering Task Force (IETF) RFC-3912).” Ex.1004, 20:22-32. A POSITA would have understood that information about the registrant such as address and contact information of the registrant could disclose the country of domain registration, since the address and contact info or the registrant is linked to the registrar (e.g., the domain name of a URL host). Further, a POSITA would have understood that the data provided by the WHOIS protocol (e.g., Internet Engineering Task Force (IETF) RFC-3912) would include address information for the registrant and may additionally include administrative and/or technical contact address information for the registrant. For example, “most whois records provide enough address information” “to be able to associate some latitude/longitude value with the IP address.” Ex.1011, 73. At a minimum, “the whois record only suggests a country or state.” *Id.* This would still be enough information to identify the claimed “country of domain registration.” For example, as of this writing, issuing a WHOIS query for “uci.edu” includes the following response, which specifies an address for the registrant, as well as an address for an administrative contact and technical contact (and a POSITA would understand WHOIS queries to include similar types of responses as of the priority date of the ‘386 Patent):

Domain Name: UCI.EDU

Registrant:

University of California, Irvine
6366 Ayala Science Library
Irvine, CA 92697-1175
USA

Administrative Contact:

Domain Admin
University of California, Irvine
6366 Ayala Science Library
Irvine, CA 92697-1175
USA
+1.9498242222
oit-nsp@uci.edu

Technical Contact:

Domain Admin
University of California, Irvine
6366 Ayala Science Library
Irvine, CA 92697-1175
USA
+1.9498242222
oit-nsp@uci.edu

Thus, a POSITA would have understood that information about the registrant such as “address and contact information of the registrant” and “registrar data including data provided by the WHOIS protocol” would include the country of domain registration for the registrant.

87. Hubbard-495 discloses or renders obvious that identifying a “web site as being based in a particular country or with a particular service provider” could be “useful in scoring because certain countries and service providers have been correlated with certain types of web content due to different laws or lax enforcement of laws.” Ex.1004, 20:8-14. It would have been obvious for a POSITA to use the “information about the registrar of the site, the identity of the registrar, the registrant (including address and contact information of the registrant)” and “any other

registrar data including data provided by the WHOIS protocol” to identify the country of domain registration and to use such information in the scoring process. In my opinion, this would qualify as the “country of domain registration” information in the claimed “the location factor of the plurality of factors for the reputation vector” comprising “two or more of: country of domain registration, country of service hosting and country of an internet protocol address block.”

88. Second, Hubbard-495 discloses or renders obvious determining a reputation score based on country of service hosting. Hubbard-495 discloses that “entire subnetworks (e.g., all addresses matching 128.2.*.* may have a particular score)” and that such “networks or subnetworks help identify a web site as ***being based in a particular country*** or with a particular service provider.” Ex.1004, 20:7-11. Hubbard-495 discloses that identifying a “web site as being based in a particular country or with a particular service provider” could be “useful in scoring because certain countries and service providers have been correlated with certain types of web content.” *Id.*, 20:8-14. It would have been obvious to a POSITA that information about a web site being based in a particular country or with a particular service provider could be combined to provide information about the country of service hosting. In addition, Hubbard-495 also discloses that “information about the URL of the web page (e.g., scoring criteria)” may include “***Geo-location data associated with the URL*** (e.g., the URL host or host domain)” and “the network or

blocks of IP addresses associated with the host IP address or its domain.” *Id.*, 18:53-19:22. A POSITA would have understood that a particular service provider provides service hosting within a particular country. A POSITA would also have understood that a service provider controls a portion of a network or subnetwork of IP addresses, which it leases out to customers to connect hosts to the internet. A POSITA would have understood that information about the country of service hosting could be derived from geo-location data associated with the URL and the service provider that provides the IP address. In my opinion, Hubbard-495’s disclosures discussed here would qualify as the “country of service hosting” information in the claimed “the location factor of the plurality of factors for the reputation vector” comprising “two or more of: country of domain registration, country of service hosting and country of an internet protocol address block.”

89. Third, Hubbard-495 discloses or renders obvious determining a reputation score based on the country of an Internet protocol address block. Hubbard-495 discloses that “information about the URL of the web page (e.g., scoring criteria)” may include “***Geo-location data associated with the URL*** (e.g., the URL host or host domain)” and “the network or ***blocks of IP addresses associated with the host IP address or its domain.***” Ex.1004, 18:53-19:22. A POSITA would have understood that information about the country of an Internet protocol address could be derived from geo-location data associated with the URL

and the name server that provides the IP address of the URL host or its domain, the network or blocks of IP addresses associated with the host IP address or its domain. In addition, Hubbard-495 also discloses “entire subnetworks (e.g., all addresses matching 128.2.*.* may have a particular score)” and that such “networks or subnetworks help identify a web site as *being based in a particular country* or with a particular service provider.” *Id.*, 20:7-11. Hubbard-495 discloses that identifying a “web site as being based in a particular country or with a particular service provider” could be “useful in scoring because certain countries and service providers have been correlated with certain types of web content.” *Id.*, 20:8-14. A POSITA would have known that “entire subnetworks (e.g., *all addresses matching 128.2.*.**)” describe internet protocol address blocks, and Hubbard-495 discloses that it would be desirable to use this information in reputation scoring. A POSITA would also be aware of top-level country codes for domain names, such as those that end in “.ru” (Russia), “.cn” (China), “.jp” (Japan), or “.dk” (Denmark),³ and Hubbard-495 discloses that in its analysis of pages and content, “the gateway server module 120 may rely at least in part on the top-level category of the domain in combination with the analysis by the reputation scoring module 331 of the content and referenced links on the website.” *Id.*, 11:20-25. In my opinion, Hubbard-495’s

³ See, e.g., https://en.wikipedia.org/wiki/Country_code_top-level_domain.

disclosures discussed here would qualify as the “country of an internet protocol address block” information in the claimed “the location factor of the plurality of factors for the reputation vector” comprising “two or more of: country of domain registration, country of service hosting and country of an internet protocol address block.” Thus, in my opinion, Hubbard-495 discloses claim 1[c3].

1[c4]: “wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and”

90. First, Hubbard-495 discloses or renders obvious determining a reputation score based on age of a domain registration. Hubbard-495 discloses that “information about the URL of the web page (e.g., scoring criteria)” may include “creation date or expiration date associated with the registration of the URL host.” Ex.1004, 18:53-19:22. Hubbard-495 further discloses that the reputation scoring module 331 may use “information about the registrar of the site (e.g., the domain name of a URL host)” such as “when the registrar record was created” to generate a reputation indicator. *Id.*, 20:22-32. It would be obvious to a POSITA to infer the age of a domain registration by simply taking the difference of the creation date associated with the registration of the URL host or the registrar record with the present date, and that such information would be desirable to use in calculating a

reputation score.⁴ A POSITA would have understood that the creation date associated with the registration of the URL host or the registrar record would thus constitute the age of a domain registration. For example, a POSITA would understand that domains with earlier creation dates (hence, larger ages) would generally be considered more stable and trustworthy. Further, Hubbard-495 discloses guarding against phishing websites “in which a HTML form or other web content is designed to appear to be provided by another, typically trusted, web site such as a bank, in order to deceive the user into providing credentials or other sensitive information to an unauthorized party.” *Id.*, 4:28-34. Of course, a POSITA would understand that a phishing website must necessarily be younger than the trusted website it is mimicking. *See, e.g., id.* In my opinion, Hubbard-495’s disclosures discussed here would qualify as the “age of a domain registration” information in the claimed “the legitimacy factor of the plurality of factors for the reputation vector” comprising “two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain.”

91. Second, Hubbard-495 discloses or renders obvious determining a reputation score based on popularity rank. Hubbard-495 discloses that “information

⁴ This is analogous to how parents with many children can keep track of their children’s ages simply by remembering their birthdates.

about the URL of the web page (e.g., scoring criteria)” may include “scoring based on site (host or domain) access statistics (e.g., generate higher reputation scores for frequently visited sites than for infrequently visited sites because frequently visited sites (for other URLs that do not have targeted content) are less likely to have targeted content in newly visited URLs).” Ex.1004, 18:53-19:22. A POSITA would have understood that scoring based on how frequently sites are visited would mean scoring based on popularity rank (i.e., popularity with visitors). In my opinion, this would qualify as the “popularity rank” information in the claimed “the legitimacy factor of the plurality of factors for the reputation vector” comprising “two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain.”

92. Third, Hubbard-495 discloses or renders obvious determining a reputation score based on internet protocol address. Hubbard-495 discloses that “information about the URL of the web page (e.g., scoring criteria)” may include “the name server that provides the IP address of the URL host or its domain, the network or blocks of IP addresses associated with the host IP address or its domain” and “historical data associated with the IP address of the URL host.” Ex.1004, 18:53-19:22. A POSITA would have understood that an “internet protocol address” factor would include scoring based on information *about* the internet protocol address, such as the name server that provides the IP address of the URL host or its

domain, the network or blocks of IP addresses associated with the host IP address or its domain, and the historical data associated with the IP address of the URL host. Hubbard-495 further discloses that the “IP address 230 may be used to identify websites that are hosting multiple domains of questionable content under the same IP address or on the same server.” *Id.*, 15:60-63. In addition, Hubbard-495 also discloses “entire subnetworks (e.g., all addresses matching 128.2.*.* may have a particular score).” *Id.*, 20:6-7. A POSITA would have known that “entire subnetworks (e.g., ***all addresses matching 128.2.*.****)” describe internet protocol address blocks, which would describe scoring based partly on IP addresses. In my opinion, this information would qualify as the “internet protocol address” information in the claimed “the legitimacy factor of the plurality of factors for the reputation vector” comprising “two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain.”

93. Fourth, Hubbard-495 discloses or renders obvious determining a reputation score based on number of hosts. A POSITA would have understood that hosts would relate to the number of IP addresses associated with the domain. The number of hosts can be inferred from the “number of IP addresses associated with the domain.” *See e.g.*, Ex.1001, Table One. For example, the ‘386 Patent discloses an exemplary trusted website with 2 hosts and an exemplary untrusted website with only 1 host. Ex.1001, 12:18-58. Consistent with this teaching from the ‘386 Patent,

a POSITA would have understood that highly-referenced (and therefore more trustworthy) domains typically have multiple IP addresses, each of which must be uniquely addressable on the Internet. For example, as of this writing, doing a DNS lookup for “netflix.com” returns the following 6 IP addresses (3 IPv4 addresses and 3 IPv6 addresses):

```
netflix.com has address 44.242.13.161
netflix.com has address 52.38.7.83
netflix.com has address 44.240.158.19
netflix.com has IPv6 address 2600:1f14:62a:de84:880a:88cc:a16:5423
netflix.com has IPv6 address 2600:1f14:62a:de83:c61a:d6e:18b0:65f7
netflix.com has IPv6 address 2600:1f14:62a:de85:d7:e7a1:8f7d:f6f5
```

A POSITA would understand that changing hosting companies also results in a change in IP address. Hubbard-495 discloses that the reputation scoring module 331 may “generate a score based on the frequency with which the IP address changes for a particular . . . domain.” Ex.1004, 20:60-64. A POSITA would have understood the frequency of IP address changes for a particular domain reflects the “number of hosts” for a particular domain within a given time period, as the frequency calculation would include the number of changes for a particular domain. Further, Hubbard-495 discloses that the reputation scoring module 331 may generate its reputation scores based on the blocks of IP addresses associated with a domain. *See, e.g.*, Ex.1004, 18:53-19:22. In my opinion, either or both of the above would qualify as the “number of hosts” information in the claimed “the legitimacy factor of the plurality of factors for the reputation vector” comprising “two or more of: age of a

domain registration, popularity rank, internet protocol address, number of hosts, top-level domain.”

94. Fifth, Hubbard-495 discloses determining a reputation score based on top-level domain. Hubbard-495 discloses “the IP address 230 may be used to identify *websites that are hosting multiple domains* of questionable content under the same IP address or on the same server.” Ex.1004, 15:61-63. Further, Hubbard-495 discloses that “the gateway server module 120 may rely at least in part on the top-level category of the domain in combination with the analysis by the reputation scoring module 331 of the content and referenced links on the website.” *Id.*, 11:21-25. In my opinion, this would qualify as the “top-level domain” information in the claimed “the legitimacy factor of the plurality of factors for the reputation vector” comprising “two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain.”

95. Thus, in my opinion, taking any two or more of the above would satisfy the limitation of “wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and.” Therefore, in my opinion, Hubbard-495 discloses or renders obvious claim 1[c4].

1[c5]: “wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency; and”

96. First, Hubbard-495 discloses or renders obvious determining a reputation score based on a plurality of run-time behaviors. A POSITA would have understood “run-time” is descriptive of “[o]ccurring after a program has begun to be executed, such as evaluation of variable expressions and dynamic allocation of memory.” *See e.g.*, Ex.1012, 460. Accordingly, a POSITA would have understood that “run-time behaviors” would refer to how a how a program might behave at run-time, including how it interacts with other files and programs. The 495 patent discloses various examples of run-time behaviors, such as writing to a registry, monitoring memory usage, processor usage etc.). Ex.1004, 21:24-28; 22:6-11. These are behaviors that would only occur once a program “after a program has begun to be executed.”

97. For example, Hubbard-495 discloses that, “after receiving candidate URLs,” the active analysis module 332 “may provide a behavioral score or data associated with one or more behavior properties (e.g., a property such as ‘writes to registry’) to the content scoring module for further categorization.” Ex.1004, 21:24-28. The active analysis module 332 may also “execute[] active content of a web page to identify its behavior properties” which may “then be used to score and classify the web page.” *Id.*, 21:16-19. A POSITA would have understood that

executing the active content of a web page to identify its behavior properties would also disclose the run-time behaviors of the web page, as this would cause the active content to begin behaving as if it had been accessed by a user. Hubbard-495 further discloses a sandbox module 370 that “transparently monitors the state of the computer to identify behavior of the web content affecting, for example, one or more of spawned processes, network access, processor usage, memory usage, use of system resources, file system access or modification, and registry access or modification.” *Id.*, 22:6-11. A POSITA would have understood these disclosures to encompass run-time behaviors of the web page/content. In my opinion, this would qualify as the “plurality of run-time behaviors” information in the claimed “behavior factor of the plurality of factors for the reputation vector” comprising “at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency.”

98. Second, Hubbard-495 discloses or renders obvious determining a reputation score based on script block counts. For instance, Hubbard-495 discloses associating categories with the candidate URLs based on the “counts of the properties for each URL,” including “keywords to be identified in the document such as ‘<javascript>.’” Ex.1004, 16:66-17:5. A POSITA would have understood that as the HTML tags, including the “javascript” tag are paired (e.g., <javascript> and

</javascript>), everything in between the paired tags represents a javascript block (or a group of javascript statements).

99. As can be seen in the Figure 13C below, Hubbard-495 teaches that HTML content that includes more than two “<javascript>” block tags may be categorized as “Malicious.”

Definitions Database

Expression	Category
Content-Type = HTML and “<javascript>” > 2	Malicious
<object> = 6 and URL matches “a.com”	Gambling
“data=,*\txt” = “data=xyx333.txt	Spyware

Ex.1004, FIG. 13C (annotated).

100. Hubbard-495 also discloses storing or renders obvious “numeric values for each property/URL corresponding to the keyword properties to indicate the number of times that the particular property was found in the web page.” Ex.1004, 17:57-61. For example: “the value 3 referring to the number of times that the property ‘<javascript>’ appears in the page.” *Id.*, 17:33-47. A POSITA would have understood that evaluation of the existence and number of times that Javascript entities are called is a type of “script” block.

101. In my opinion, these disclosures would qualify as the “script block count” information in the claimed “behavior factor of the plurality of factors for the

reputation vector” comprising “at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency.”

102. Third, Hubbard-495 discloses or renders obvious determining a reputation score based on picture count. Hubbard-495 teaches “potentially executable web content may also be malicious content” and that “image files have been used to exploit vulnerabilities in certain operating systems when those images are processed for display.” *Id.*, 4:25-28. Similar to the script block tag discussed above, a POSITA would have understood that image files may be embedded in web pages using an “” tag. Hubbard-495 also discloses storing “numeric values for each property/URL corresponding to the keyword properties to indicate the number of times that the particular property was found in the web page.” Ex.1004, 17:57-61. “The content filtering module 150 may be configured to parse the returned page and identify URL links in the returned search results and embedded items of content such as images or active content.” *Id.*, 8:37-41. A POSITA would have understood that the “particular property” “found in the web page” would include the items parsed by the content filtering module, such as “embedded items of content such as images or active content.” A POSITA would have understood that pictures would constitute “content such as images.” In my opinion, these disclosures would qualify as the “picture count” information in the claimed “behavior factor of the plurality of factors for the reputation vector” comprising “at least one of: plurality

of run-time behaviors, script block count, picture count, immediate redirect and response latency.”

103. Thus, in my opinion, taking any one or more of the above, Hubbard-495 discloses “wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency; and.” Therefore, in my opinion, Hubbard-495 discloses or renders obvious claim 1[c5].

1[d]: “transmitting a decision transmission to the Internet-enabled client application of the client-side device, the decision transmission allowing or denying access to the Internet resource.”

104. As discussed above with respect to element 1[a], an Internet-enabled client application of the client-side device of a local area network transmits a request for an Internet resource, and a POSITA would have understood that the decision transmission would be a *response* to the request. Hubbard-495 discloses client computers that may include “software modules such as *web browsers*,” and that each of the client computers may be in “communication with a gateway server module 120,” which is “in two way communication” with the client computers and the Internet. Ex.1004, 4:46-51, 6:19-23. “Thus, *requests* originating within” the client computers “may be required to pass through the gateway server module 120 as they proceed to the Internet.” *Id.*, 6:24-26. Hubbard-495 further discloses that the URL request from the client computer may be “intercepted at the Internet gateway and

forwarded to the gateway server module 120,” and the URL record is analyzed “to determine whether the category of the URL is one that should be blocked for the workstation user.” *Id.*, 22:53-62. “If the category is blocked, the process skips to block 1212 and the request is blocked. If the category is not blocked, however, the request is allowed at block 1208.” *Id.*, 22:62-65.

105. Further, Hubbard-495 discloses or renders obvious a method of controlling access to web content, which “further comprises determining whether to ***allow the request*** based at least partly on the at least indicator.” Ex.1004, 2:33-40. Hubbard-495 discloses that the gateway server module 120 “determines whether to ***allow access*** to the URL based at least partly on the reputation score.” *Id.*, 27:51-53. The gateway server module 120 may also “rely upon the determined category in combination with the score associated with the URL when determining whether to ***allow the user access*** to the website associated with the requested URL.” *Id.*, 27:58-61. A POSITA would have understood that “determining whether to allow access” would necessarily encompass determination outcomes of both allow and disallow (*i.e.*, deny). A POSITA would understand that in the system of Hubbard-495 a decision to allow access to an Internet resource is necessarily communicated to the client device (workstation 116) in that the allowed web content is delivered to the workstation 116. It would have been obvious to a POSITA to transmit that responsive decision transmission to the Internet-enabled client application of the

client-side device that made the request in the first place. A POSITA would have understood that an end user would not be able to view the allowed request unless the response had been transmitted *back* to the Internet-enabled client application of the client-side device that made the request

106. For example, the ‘386 Patent describes transmitting a “decision transmission” to the client that differs depending on whether access to the Internet resource was allowed or not. In a situation in which access is allowed, the ‘386 Patent explains “the decision transmission is the Web page for the Web site.” Ex.1001, 5:37-42.

107. Further, Hubbard-495 discloses that it is concerned with the latency of responses to web content requests. *See, e.g.*, Ex.1004, 9:34-38. A POSITA would have understood that it would be desirable to transmit a decision transmission to the Internet-enabled client application on the workstations, for example, if the web content is blocked. Otherwise, a user may be confused that the request for the web content has merely been lost or delayed. A POSITA would be aware that absent a transmission of the decision to the Internet-enabled client application of the client-side device in the case of denying access to the Internet resource, there would be no way for a client device to distinguish between a website that is taking a long time to be returned and one that has been blocked. Thus, a POSITA would understand that in the Hubbard-495 system it is inherent or at least suggested that a transmission is

returned to the client in the case of blocked web content. *See, e.g., Id.*, Figs. 16, 17, 18, 25, 22:53-65, 23:4-10, 23:35-39, 23:48-52, 23:60-64. Further, Hubbard-495 discloses that “[t]he gateway server module 120 is in two way communication with the workstation 116.” *Id.*, 6:19-20.

108. Hubbard-495 discloses that the gateway server module 120 may also include a content filtering module 150 that filters “portions of search results returned in response to a network request.” Ex.1004, 8:24-28. “A page returned by the search engine 113 to the client workstation 116 may be intercepted by the gateway server module 120 and analyzed by the content filtering module 150.” *Id.*, 8:34-37. “The content filtering module 150 may be configured to parse the returned page and identify URL links in the returned search results and embedded items of content such as images or active content,” and may also “remove URL links and other content from a list of URLs returned in response to a user query sent to an Internet search engine 113.” *Id.*, 8:28-41. A POSITA would have understood that allowance or denial of user access would be transmitted to the client or user as a response to the “user query.” Thus, in my opinion, Hubbard-495 discloses or renders obvious claim 1[d].

109. In addition, this claim limitation is obvious in view of other work by one of Hubbard-495’s inventors. Hubbard-174 “specifies the action taken with respect to each data item such as ‘Observe’, ‘*Allow*’, or ‘*Block*.’” Ex.1005, 13:20-

22; *see also id.*, Fig. 11B. A POSITA would understand that in the system of Hubbard-174 a decision to allow access to an Internet resource is necessarily communicated to the client device (workstation 116) in that the allowed web content is delivered to the workstation 116.

250

ITEM TYPE	URL	ACTION TAKEN	CATEGORY	CODE
a,href	http://www.coolscreensavers.com/	Allow	Search Engines	76
a,href	http://www.verycoolscreensavers.com/	Allow	Search Engines	76
a,href	http://www.verycoolscreensaver.com/	Observe	Free Software	99
a,href	http://www.somewhatcoolscreensavers.com/	Allow	Search Engines	76
a,href	http://www.downloadacoolscreensaver.com/	Observe	Free Software	99
img,src	http://www.adultcoolscreensavers.com/sample.gif	Block	Adult Content	66
a,href	http://www.thebestscreensavers.com/	Allow	Search Engines	76
a,href	http://www.artisticscreensavers.com/	Allow	Search Engines	76
img,src	http://www.verycoolscreensavers.com/save.jpg	Allow	Search Engines	76
a,href	http://www.trickyscreensavers.com/	Block	Spyware	154
a,href	http://www.basicscreensavers.com/	Observe	Free Software	99

252 254 256 258 260

FIG. 11B

Ex.1005, Fig. 11B.

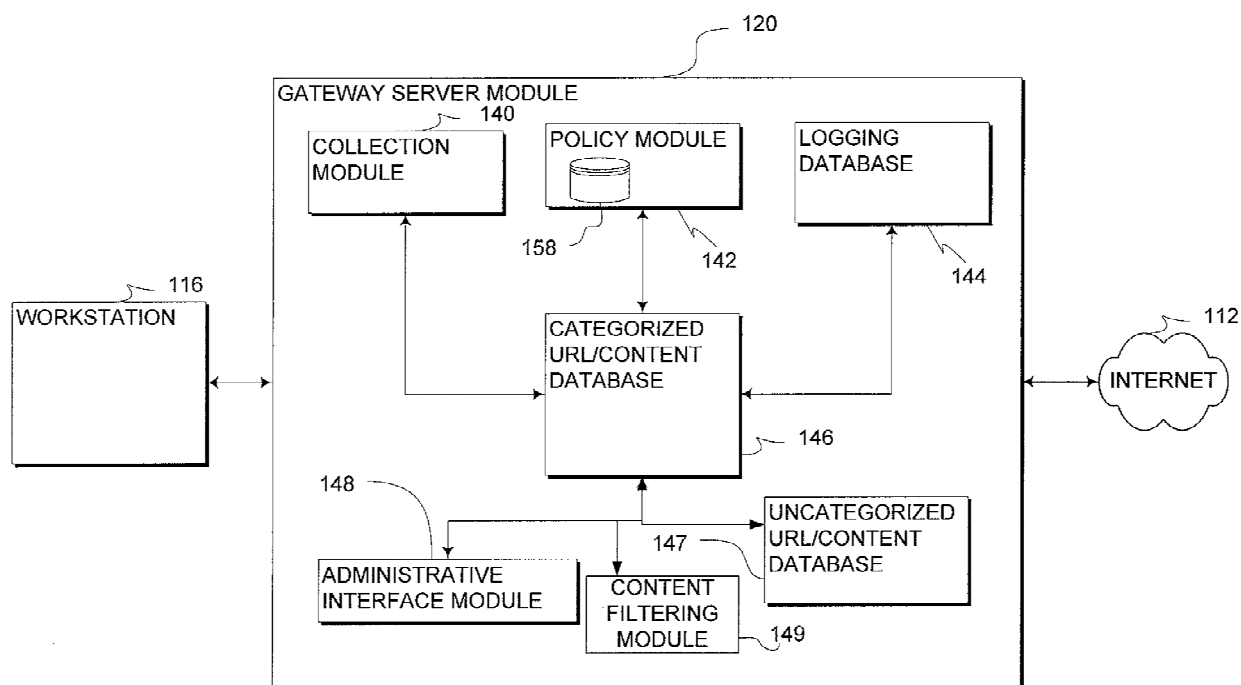


FIG. 3

Ex.1005, Fig. 3.

110. Like Hubbard-495 (and the ‘386 Patent), Hubbard-174 discloses a “content filtering module” that modifies Internet content to only show users material that is deemed safe. Ex.1005, 6:49-65. Hubbard-174 discloses that “the *modifications* to the image search page 280 performed by the content filtering module 149 *are visible to the user and shown as blocks of disallowed data*” while the page 280 “includes *allowed images 282 returned by the search engine 113 and unobstructed by the gateway server module 120*, and also includes notifications of *blocked images 284.*” *Id.*, 14:46-52. The blocked images 284 may also be “replaced

by a display element *notifying the user that the image has been blocked.*” *Id.*, 14:52-54. A POSITA would have understood that in order for the user to be notified or for the modifications to be visible to the user, the page 280 would need to be transmitted to the Internet-enabled client application of the client-side device, e.g., as shown below.

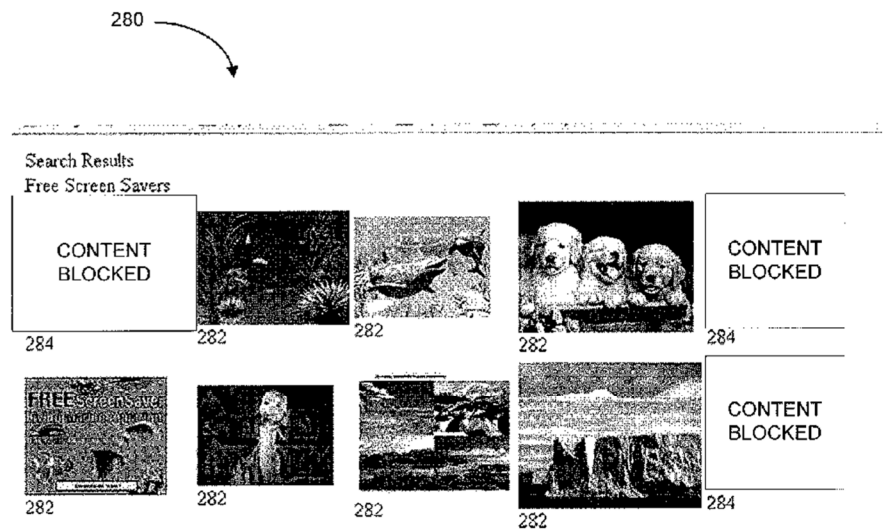


FIG. 14

Ex.1005, Fig. 14.

111. As shown in the figure above, the allowed images constitute decisions to allow access to an Internet resource and are transmitted to the user (as shown by the images appearing on the display). The blocked images constitute decisions to deny access to an Internet resource, and the notifications that content has been

blocked are transmitted to the user (as shown by the “CONTENT BLOCKED” boxes on the display). A POSITA would have understood that, in this situation, Hubbard-174 transmitted a decision transmission to the Internet-enabled client application of the client-side device, the decision transmission allowing or denying access to the Internet resource. Such application of a decision transmission would have yielded nothing more than the predictable results of enhancing the system of Hubbard-495 by adding the functionalities of Hubbard-174 without the need for rearrangement or excessive modification of the system of Hubbard-495.

Motivation to Combine Hubbard-495 and Hubbard-174

112. As I explain above, and include by reference here, in my opinion, a POSITA would have been motivated to combine Hubbard-495 and Hubbard-174. For example, as I explain above, both are analogous art to the ‘386 Patent at least because each is from the same field of endeavor, namely, methods and systems to determine conditions under which access to an Internet resource is allowed. Ex.1004, 2:33-61; Ex.1005, 1:7-9. A POSITA at the relevant time would have understood that Hubbard-495 and Hubbard-174 could be readily combined as they disclose the same exemplary system, and they have an inventor (Hubbard) in common. Ex.1004, cover, 4:38-5:7; Ex.1005, cover, 4:16-52. For example, both disclose embodiments that include a gateway server module connected through the Internet to a database management module. Ex.1004, 4:50-67; Ex.1005, 4:29-46.

In those embodiments, workstations make requests for web content that may be allowed or blocked by the gateway server module. *See* Ex.1004, 6:34-36, 6:53-7:3; Ex.1005, 6:10-12, 6:27-45. It would be desirable to transmit a decision transmission to the Internet-enabled client application on the workstations, for example, if the web content is blocked. A POSITA would be aware that absent a transmission of the decision to the Internet-enabled client application of the client-side device in the case of denying access to the Internet resource, there would be no way for a client device to distinguish between a website that is taking a long time to be returned and one that has been blocked. Such application of a decision transmission would have yielded nothing more than the predictable results of enhancing the system of Hubbard-495 by adding the functionalities of Hubbard-174 without the need for rearrangement or excessive modification of the system of Hubbard-495.

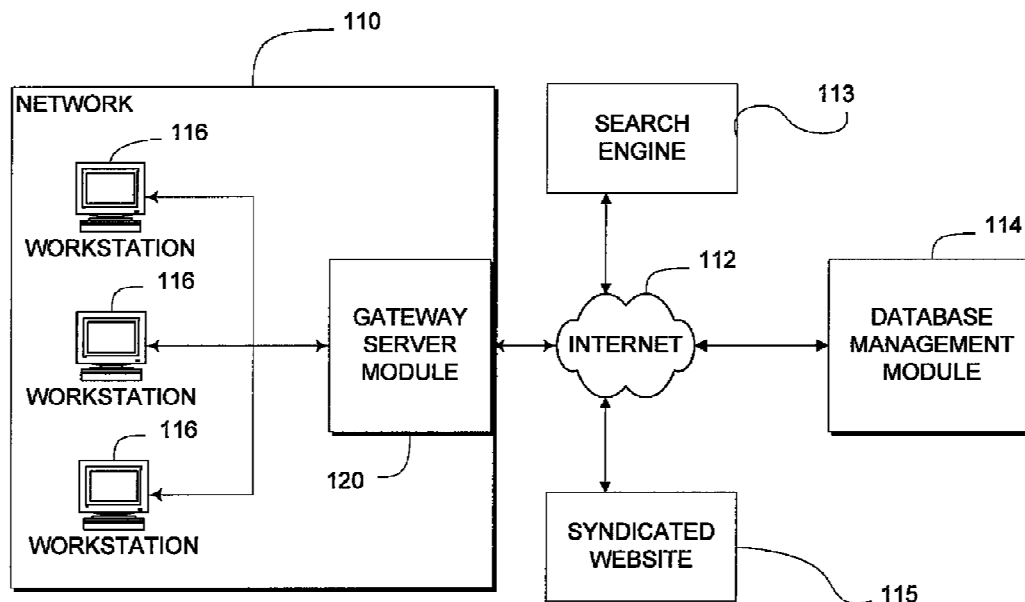


FIG. 1

Ex.1004, Fig. 1; Ex.1005, Fig. 1.

B. Claim 2

“The method according to claim 1, wherein the reputation index for the Internet resource is at or above the threshold value and wherein the decision transmission is access to the Internet resource, the method further comprises transmitting the request for the Internet resource over the Internet to the first server for the Internet resource and receiving the Internet resource at the local area network.”

113. In my opinion, Hubbard-495, either alone or in combination with Hubbard-174, discloses or renders obvious this limitation. As I discussed above with respect to element 1[c], Hubbard-495 describes that the URL analysis module

149 may generate a “**sufficiently high** reputation score (e.g., indicative of a good reputation or low likelihood of active or other targeted content) for an uncategorized URL so that the policy module 142 **allows access to the uncategorized URL** when the policy module 142 would otherwise not allow access to the URL.” Ex.1004, 11:29-36. A POSITA would have understood that determining whether a reputation score is “sufficiently high” would involve comparing it to a threshold value. Hubbard-495 also contemplates the use of thresholds in the scoring process, where, e.g., “categories corresponding to each score above a given **threshold** are identified with the URL.” *Id.*, 21:1-4. *See also, e.g., id.*, 19:45-51, 19:55-61, 20:65-21:1 (“In one embodiment, **URLs with scores above a threshold** are identified with a category, e.g., malicious, regardless of, or in addition to, the category identified by content analysis.”), 21:1-10. Hubbard-495 thus discloses the claimed “the reputation index for the Internet resource is at or above the threshold value and wherein the decision transmission is access to the Internet resource.”

114. As I discussed above with respect to elements 1[a] and 1[b], and for the same reasons, Hubbard-495 discloses or renders obvious “transmitting the request for the Internet resource over the Internet to the first server for the Internet resource.”

115. As I discussed above with respect to element 1[d], and for the same reasons, Hubbard-495, alone or in combination with Hubbard-174, discloses or

renders obvious a “decision transmission allowing access to the Internet resource.” A POSITA would have understood that “allowing access” would mean allowing the client device to receive the Internet resource requested at the client device’s local area network 110. For example, Hubbard-174 describes that “[w]hen a user submits a search request to the search engine 113,” “results are returned to the user in a web page which displays data about the relevant pages and URL links which allow the user to access the identified pages.” Ex.1005, 12:30-39. Hubbard-495 discloses that the gateway server module 120 may also include a content filtering module 150 that filters “portions of search results returned in response to a network request.” Ex.1004, 8:24-28. A POSITA would have understood that such search results “returned” in response to a request would constitute receiving the Internet resource at the client device of the local area network.

116. As such, for the same reasons discussed above with respect to elements 1[a], 1[b], 1[c], and 1[d], Hubbard-495, either alone or in combination with Hubbard-174, also discloses or renders obvious claim 2.

C. Claim 3

“The method according to claim 1, wherein the reputation index for the Internet resource is below the threshold value, and wherein the decision transmission is a Web page from the local area network containing information that accessing the Internet resource is against an Internet policy established on the local area network.”

117. In my opinion, Hubbard-495, either alone or in combination with Hubbard-174, discloses or renders obvious this limitation. As I discussed above with respect to element 1[d], and for the same reasons, Hubbard-495 in combination with Hubbard-174 discloses or renders obvious a “decision transmission allowing or denying access to the Internet resource.”

118. As I discussed above with respect to element 1[c], Hubbard-495 describes that the policy module 142 “may be configured to block access to URLs with low scores (e.g., indicative of poor reputation or higher likelihood of active or other targeted content).” Ex.1004, 11:36-40. The policy module 142 may be “used to *implement network policies* regarding how certain content will be handled by the gateway server module 120 or by a firewall or some other security software installed within the network 110.” *Id.*, 6:54-58. A POSITA would have understood a network policy to mean a policy established for the network 110, which may be a local area network as I discussed above with respect to element 1[a].

119. The policy module 142 “may be configured to *block access to URLs with low scores* (e.g., indicative of poor reputation or higher likelihood of active or other targeted content).” Ex.1004, 11:36-40. A POSITA would understand “low scores” means the reputation index for the Internet resource is below the threshold value. It would have been obvious to a POSITA that if access to the URL was blocked, that the Internet resource should not be received at the network 110, and

that the decision transmission to the Internet-enabled client application of the client-side device would not include the Internet resource. A POSITA would have understood that in the event that the decision transmission would not be the Internet resource in the case of blocked access, the decision transmission would still have to be a Web page from the local area network. As discussed above with respect to elements 1[a] and 1[d], the decision transmission is in response to a request from an Internet-enabled client application of the client-side device and is sent *back* to the Internet-enabled client application of the client-side device. The most natural and logical approach would be to send the response in a format readable by the Internet-enabled client application, which may be the same format as what was requested, *i.e.*, a Web page.

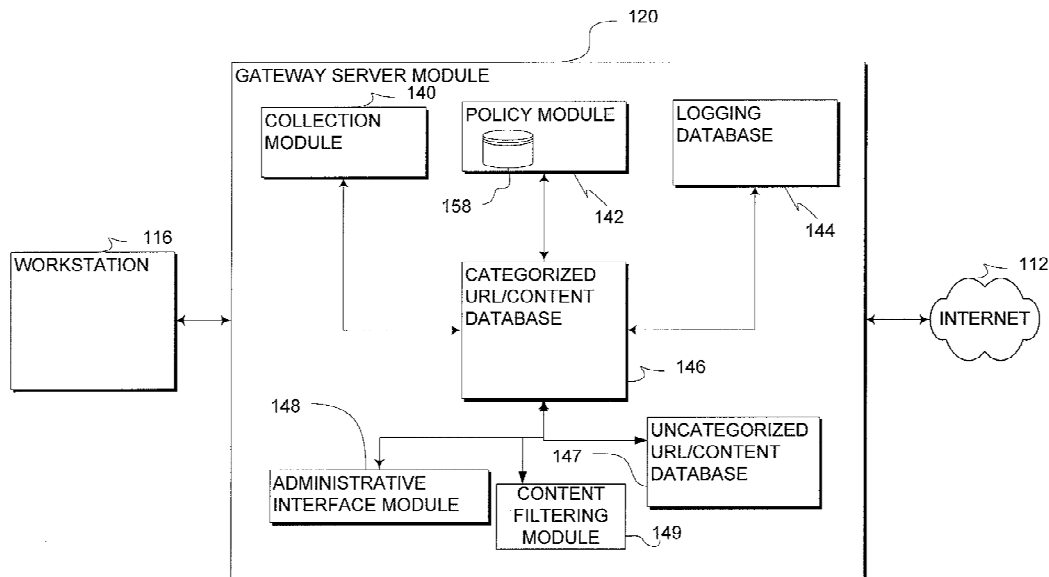
120. In at least one embodiment of Hubbard-495, “the ***gateway server module 120 determines whether to allow access to the URL*** based at least partly on the reputation score.” Ex.1004, 27:51-53. As discussed above with respect to element 1[b], the gateway server module 120 is part of the local area network 110. A POSITA would have understood that it would be advantageous for a variety of reasons, such as performance, privacy, and security factors, to have the decision transmission remain within the local area network.

121. As discussed above with respect to element 1[d], Hubbard-495, alone or in combination with Hubbard-174, discloses or renders obvious a “decision

transmission” allowing or denying access to the Internet resource. It would have been obvious to a POSITA to relay the decision transmission to the Internet-enabled client application of the client-side device in the case of denying access to the Internet resource, in order to distinguish between a website that is taking a long time to be returned and one that has been blocked. It would further have been obvious to a POSITA that it would be desirable to include information that accessing the Internet resource is against an Internet policy, so that the user would not attempt to access the blocked URL again.

122. Further, this claim limitation is obvious in view of Hubbard-174, which discloses that the “gateway module identifies an outgoing network request” and that “[w]hen data is sent back to the workstation 116 in response to the network request, the gateway server module 120 intercepts the returned data.” Ex.1005, 15:50-57. The content filtering module 149 of the gateway server module 120, *see id.*, Fig. 3, “modifies the content by either ***removing or blocking the objectionable content*** as identified in accordance with the policies defined in the policy module 142.” *Id.*, 16:21-25. “Once the content filtering module 148 has completed its modifications of the intercepted page, ***the modified page*** is then sent to the originating workstation 116.” *Id.*, 16:25-28. A POSITA would have understood that this modified page, created at the gateway server module 120 of the local area network, would necessarily be a Web page from the local area network. A POSITA would have

understood that since the gateway server module 120 creates the modified page, it would inherently be a Web page created at the local area network.



Ex.1005, Fig. 3.

123. Hubbard-174 discloses that “the *modifications to the image search page 280 performed by the content filtering module 149 are visible to the user* and shown as blocks of disallowed data” while the page 280 “includes allowed images 282 returned by the search engine 113 and unobstructed by the gateway server module 120, and *also includes notifications of blocked images 284.*” Ex.1005, 14:46-52. A POSITA would have understood that in order for the user to view the image search page 280, the page 280 would need to be transmitted to the Internet-enabled client application of the client-side device. The blocked images 284 may also be “replaced by a display element *notifying the user that the image has been*

blocked.” *Id.*, 14:52-54. “For example, the notification may indicate that the URL link would access a page categorized as adult content in the URL/ content database 146 and that the policy module 142 has been configured to block such content.” *Id.*, 13:63-67. A POSITA would have understood that the modified page 280 returned to the user with a display element notifying the user of that content has been blocked would constitute “a Web page from the local area network containing information that accessing the Internet resource is against an Internet policy established on the local area network.”

D. Claim 4

“The method according to claim 1, wherein the Internet resource is one of a Web site, an email server and a voice communications server.”

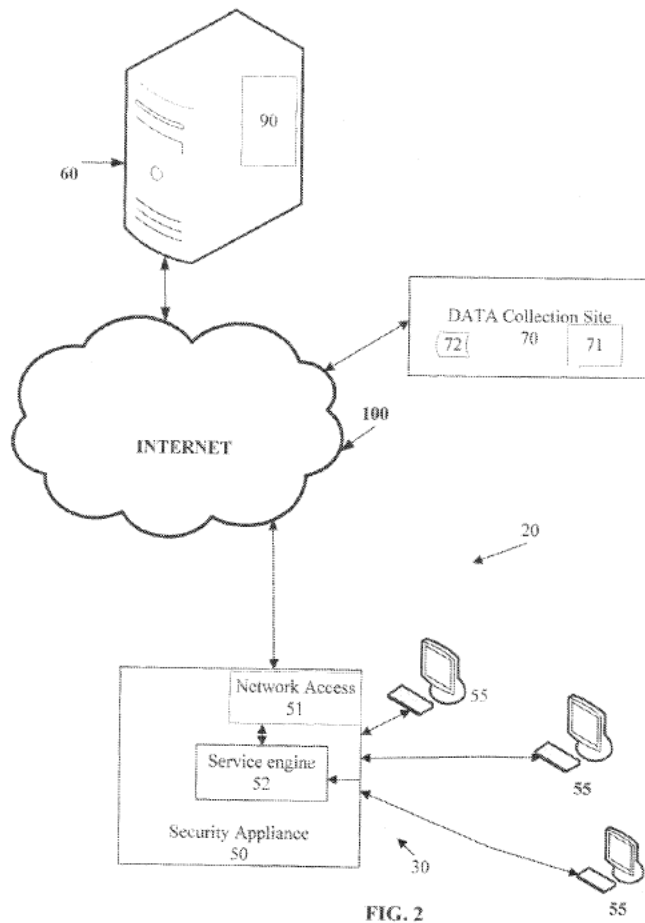
124. In my opinion, Hubbard-495 discloses or renders obvious this limitation. Hubbard-495 discloses that the client computers 116 “may also be loaded with operating systems that allow them to utilize the network through various software modules such as *web browsers, e-mail programs, or the like.*” Ex.1004, 4:46-49. It would be obvious to a POSITA that web browsers are set up for the purpose of accessing Web sites and that e-mail programs are configured to retrieve emails from an email server. A POSITA would have also understood that the Internet resource could also be a voice communications server, as the use of VoIP calls was well understood at the time.

E. Claim 5

“The method according to claim 1, wherein the reputation index is received from a data collection site accessible by the security appliance over the Internet.

125. In my opinion, Hubbard-495 discloses or renders obvious this limitation. As I discussed above with respect to element 1[c], Hubbard-495 discloses a score representative of a reputation of the URL, which qualifies as the claimed “reputation index.” As I discussed above in element 1[b], Hubbard-495 discloses or renders obvious a gateway server module 120, which qualifies as the claimed “security appliance.”

126. The ‘386 Patent shows that the data collection site 70 is connected to a security appliance via the internet. Ex.1001, FIG. 2. The ‘386 Patent describes a reputation generating site 70, “preferably having a crawler 71 and a database 72”, that “provides reputation indices to service engine 52 of the security appliance 50.” *Id.*, 10:42-58.



Ex.1001, Fig. 2.

127. In my opinion, a POSITA would have understood that the database management module 114 of Hubbard-495 qualifies as the claimed “data collection site.” As seen in Hubbard-495 FIG. 1, the database management module 114 is connected to the gateway server module 120 over the Internet. As seen in FIG. 7, the database management module 114 may also include a URL/content database 180 and a collection system 182.

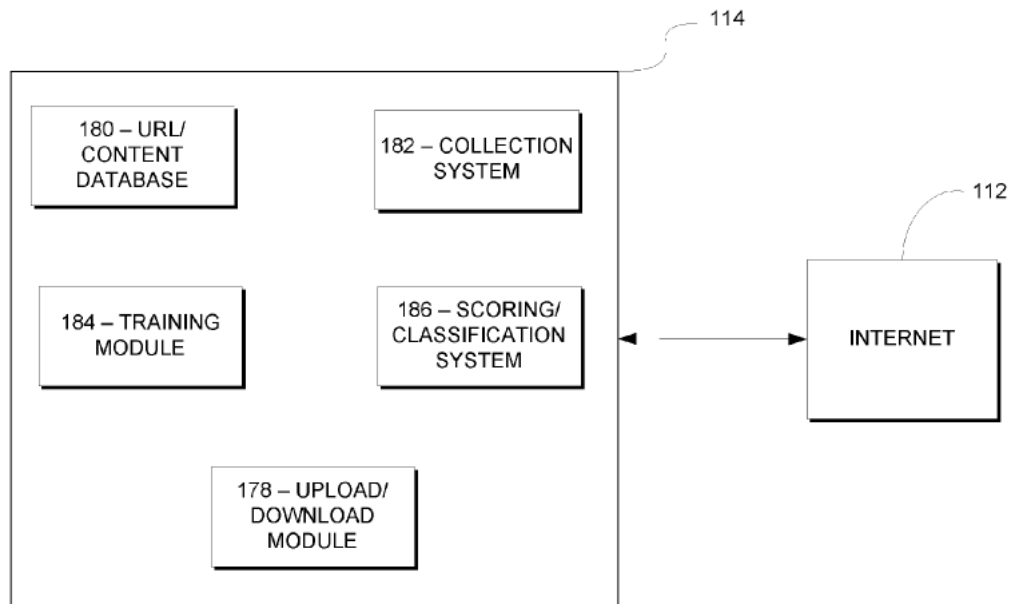


FIG. 7

Ex.1004, Fig. 7.

128. The collection system 182 “may include a database management collection module 190” which may include various subsystems such as the webcrawler module 206 and the honey client module 208. Ex.1004, 12:55-58; 13:59-14:17. Thus, Hubbard-495 discloses or renders obvious the database management module 114 which is connected to the gateway server module 120 over the Internet and includes a database and a web crawler.

129. Hubbard-495 discloses that the URL analysis module 149 “may rely upon a remotely located reputation scoring module,” which may be a reputation scoring module 331 “located at the database management module 114.” Ex.1004,

9:43-49. The URL analysis module 149 “communicates the requested URL to the database management module 114 *for scoring by the remotely located reputation scoring module 331*” and the database management module 114 “communicates the results of the analysis performed by the reputation scoring module 331 *back to the gateway sever module 120.*” *Id.*, 9:49-54. A POSITA would have understood that the “results of the analysis performed by the reputation scoring module 331” would have been the results of “scoring by the remotely located reputation scoring module 331.” Thus, Hubbard-495 discloses that the remotely located reputation scoring module 331 of the database management module 114 generates reputation scores and sends them back to the gateway server module. In my opinion, this qualifies as the data collection site 70 of the ‘386 Patent that “provides reputation indices to service engine 52 of the security appliance 50.”

130. A POSITA would have understood that the sending of scores by the database management module 114 to the gateway server module 120 would encompass receiving the scores at the gateway server module 120 from the database management module 114. In my opinion, this would qualify as the claimed “wherein the reputation index is received from a data collection site accessible by the security appliance over the Internet.”

F. Claim 6

“The method according to claim 1, wherein the reputation index is stored at a remote reputation generating site.”

131. In my opinion, Hubbard-495 discloses or renders obvious this limitation. It would have been obvious to a POSITA that “remote” would mean a device that is not on the local area network. Hubbard-495 discloses that the URL analysis module 149 “may rely upon a remotely located reputation scoring module,” which may be a reputation scoring module 331 “located at the database management module 114.” Ex.1004, 9:43-49. The database management module 114 “may be located remotely (accessible via Internet 112) from the network 110 and its associated workstations 116.” *Id.*, 12:5-8. As I discussed above with respect to Claim 5, the database management module 114 generates a score for a requested URL. In my opinion, the database management module would qualify as the claimed “remote reputation generating site.” A POSITA would have understood that in generating a score, the database management module 114 would necessarily have to store said generated score, such as transiently in memory.

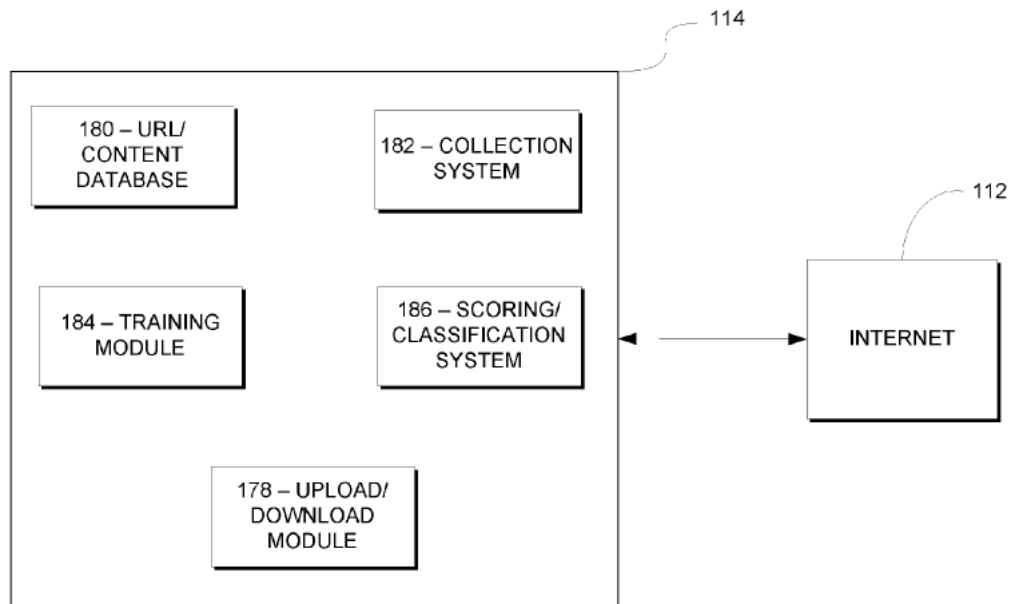


FIG. 7

Hubbard-495, Fig. 7.

132. As seen in FIG. 7, “the database management module 114 may also include a database management module URL/content database 180.” Ex.1004, 12:25-26. Hubbard-495 contemplates storing information about the URL that is used to make a determination about whether to allow or block access. “[T]he static content classification module 328 *stores* the categories of each URL to a categorized web page database. In one embodiment, the URL database 180 includes the categorized web page database.” *Id.*, 27:5-9. Hubbard-495 further discloses that “[o]nce *the data has been collected and stored*, the process moves to block 1410

where the URL request is either allowed or blocked based on the policies indicated in the policy module 142.” *Id.*, 23:60-64.

133. Hubbard-495 also teaches that the “database management module URL/content database 180 may be a relational database that is indexed to provide quick and effective searches for data.” Ex.1004, 12:31-33. Further, Hubbard-495 discloses that “[a]dvantageously, a remotely located reputation scoring module may be shared by multiple gateway server modules 120.” *Id.*, 9:61-63. It would have been obvious to a POSITA to store a reputation score in the database to “provide quick and effective searches for data” and to minimize repeated calculations if a reputation score is requested for the same URL. A POSITA would have understood that the reputation score could be stored in the URL/content database of the remotely located database management module 114 for greater efficiency and information dispersal across multiple gateway server modules 120, and that it would have been obvious to do so. Further, it would have been obvious to a POSITA that storing the reputation score at a remote reputation generating site —such as at the remotely located database management module 114 of Hubbard-495—would be desirable for data backups and mirroring purposes, which were common techniques in software design at the time.

G. Independent Claim 7

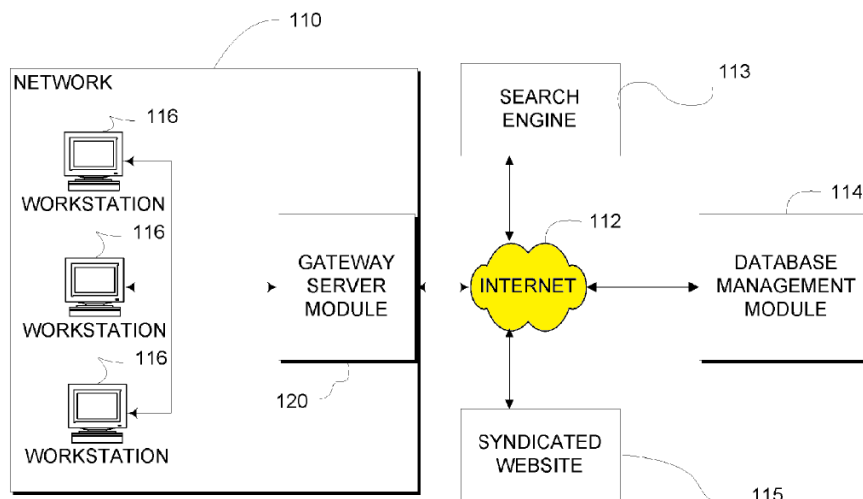
134. In my opinion, Hubbard-495, alone or in view of Hubbard-174, discloses or renders obvious claim 7.

7[preamble]: “A system for controlling access to a Web site, the system comprising:”

135. In my opinion, Hubbard-495 discloses or renders obvious the preamble. Hubbard-495 discloses “a system for controlling access to web content.” Ex.1004, 2:41-42; 2:52-53. It would have been understood by a POSITA that “web content” would constitute a “Internet resource,” such as a “Web site.”

7[a]: “a network, the network comprising the Internet;”

136. Hubbard-495 discloses or renders obvious this limitation. For example, Hubbard-495 teaches that the “gateway server module may reside at the edge of the network 110 so that traffic sent to and from the Internet 112 may pass through it on its way into or out of the network 110.” Ex.1004, 4:51-54. A POSITA would have understood that the Internet is inherently a network.

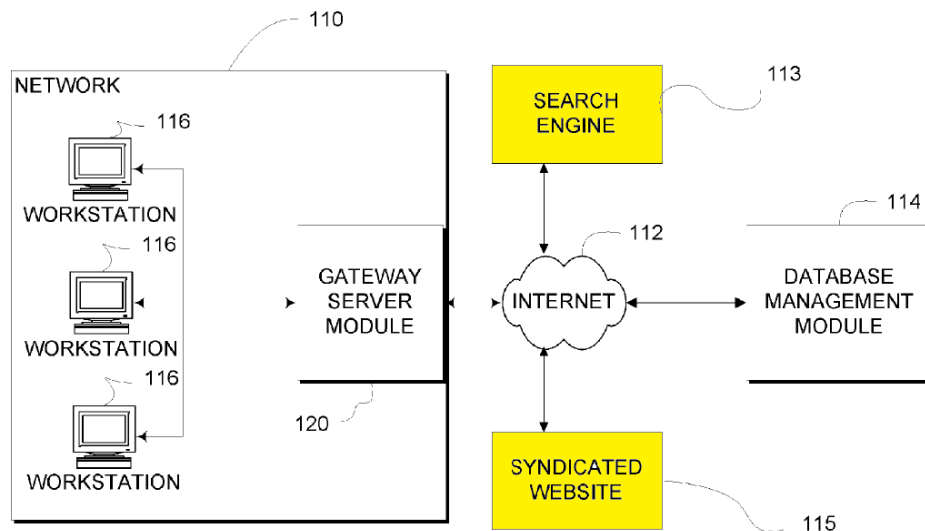


Ex.1004, Fig. 1 (annotated).

7[b]: “an Internet resource hosted at a first server and accessible over the Internet; and”

137. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[a]. As explained with respect to 1[a] above, the workstations 116 may be in communication with “*one or more search engines 113 via the Internet 112.*” Ex.1004, 5:1-7. The search engine 113 may “return one website or a list of websites (and URL links leading to those websites)” in response to the search. *Id.*, 5:45-47. Alternatively, the request may also be “a content request made to the syndicated website 115” which may return “one URL link or a list of URL links to the requested content.” *Id.*, 5:48-51. A POSITA would have understood that a “website or a list of websites (and URL links leading to those websites)” and “URL link or a list of URL links to the requested

content” would be examples of webpages. It was well known that webpages would be hosted at a web server. A POSITA would have understood that webpages on a web server would be an “Internet resource hosted at a first server.” This understanding is likewise confirmed in Hubbard-495, e.g., *id.* 5:8-11 (“In the example system of FIG. 1, one or more of the workstations 116 may be in communication with one or more syndicated websites 115 via the Internet 112. The syndicated website 115 offers a web feed of content from the website.”). A POSITA would have understood that the search engine 113 and the syndicated website 115 are examples of Internet resources that would be hosted at a server.



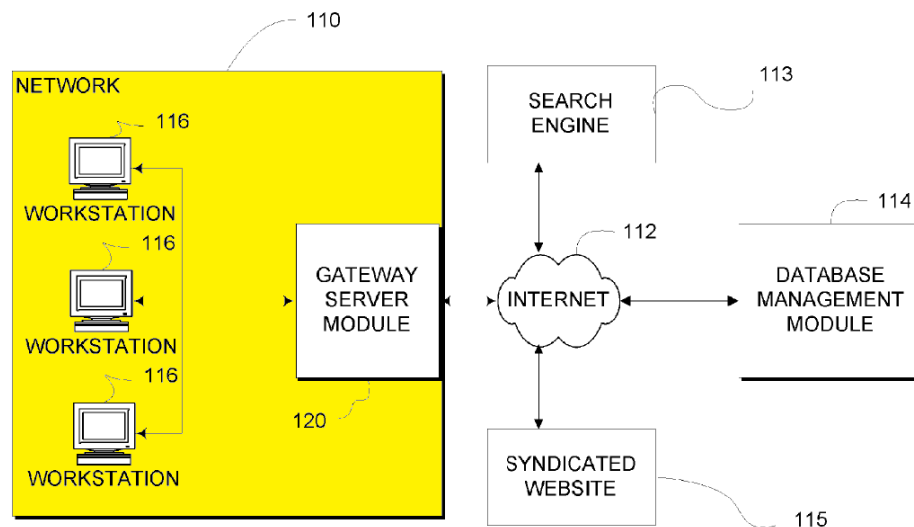
Ex.1004, Fig. 1 (annotated).

138. In my opinion, the search result content on a search engine 113 and/or website content on a syndicated website 115 (being requested by the workstation

116) would qualify as the “Internet resource residing at a first server.” Thus, in my opinion, Hubbard-495 discloses claim 7[b].

7[c]: “a local area network, the local area network comprising:”

139. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[a].

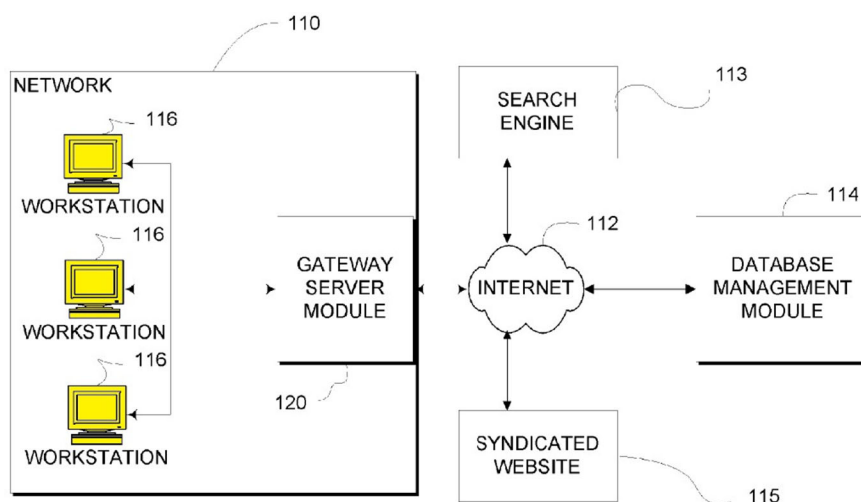


Ex.1004, Fig. 1 (annotated).

7[d]: “(a) a plurality of client-side devices, each of the client-side devices having a browser; and”

140. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[a]. Hubbard-495 discloses an exemplary system that includes a network 110, which may “be a *local area network*,” and “may include one or more workstations 116” which may be “*various types of client computers that are attached to the network.*” Ex.1004,

4:39-44. The client computers may “be loaded with operating systems that allow them to utilize the network through various software modules such as *web browsers*, e-mail programs, or the like.” *Id.*, 4:46-49. In my opinion, this would qualify as the claimed “plurality of client-side devices, each of the client-side devices having a browser.” Thus, in my opinion, Hubbard-495 discloses claim 7[d].



Ex.1004, Fig. 1 (annotated).

7[e]: “(b) a security appliance, the security appliance controlling access to the Internet by each of the plurality of client-side devices, the security appliance having a service engine for analyzing a reputation vector for the Web site and generating a reputation index for the Web site from the reputation vector, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location, wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block, wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more

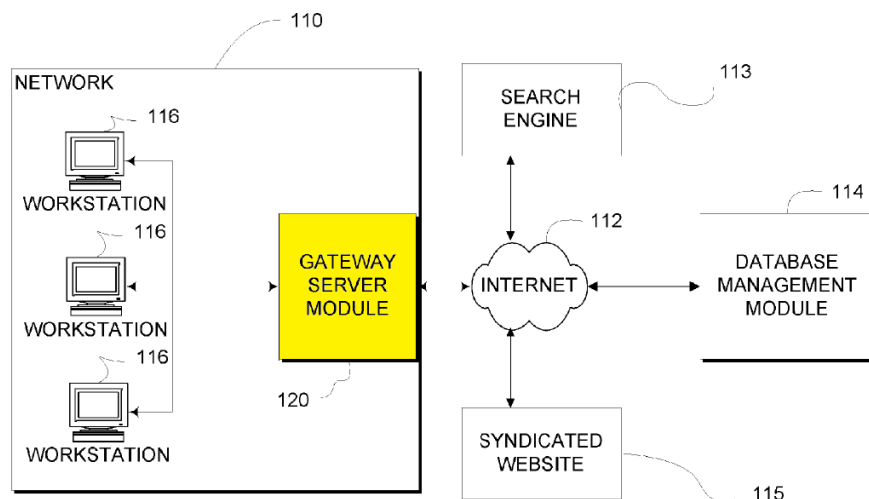
of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency;”

141. In my opinion, Hubbard-495 discloses or renders obvious this limitation, as I explain below.

7[e1]: “(b) a security appliance, the security appliance controlling access to the Internet by each of the plurality of client-side devices,

142. As I discussed above with respect to element 7[d], and for the same reasons, Hubbard-495 discloses or renders obvious a “plurality of client-side devices.”

143. As I discussed above with respect to element 1[b], the gateway server module 120 of Hubbard-495 would constitute the claimed “security appliance.”



Ex.1004, Fig.1 (annotated).

144. Hubbard-495 discloses that “*requests originating within the workstations 116 of the network 110 may be required to pass through the gateway server module 120 as they proceed to the Internet.*” Ex.1004, 6:24-26. The gateway server module 120 “may receive URL requests” from the workstations 116. *Id.*, 6:34-36. Hubbard-495 also discloses that “the URL analysis module 149 of the gateway server module 120 of FIG. 1 receives a request for a URL from a workstation 116.” *Id.*, 27:34-37. Hubbard-495 further discloses that the gateway server module 120 “determines whether to *allow access* to the URL based at least partly on the reputation score.” *Id.*, 27:51-53. The gateway server module 120 may also “rely upon the determined category in combination with the score associated with the URL when determining whether to *allow the user access* to the website associated with the requested URL.” *Id.*, 27:58-61. A POSITA would have understood that by determining whether to allow access to the website and/or URL, the gateway server module 120 controls access to the Internet. Thus, in my opinion, Hubbard-495 discloses or renders obvious claim 7[e1].

7[e2]: “the security appliance having a service engine for analyzing a reputation vector for the Web site and generating a reputation index for the Web site from the reputation vector,

145. Hubbard-495 discloses that after “the URL analysis module 149 of the gateway server module 120” “receives a request for a URL from a workstation 116,” “a reputation scoring module generates *a reputation score* or other reputation

indicator of targeted content associated with the URL.” Ex.1004, 27:34-40. “The reputation scoring module *may be located locally at the gateway server module 120.*” *Id.*, 27:40-42. “When located within the gateway server module 120, the reputation scoring module may *directly analyze* the reputation of the URL.” *Id.*, 9:34-38. A POSITA would understand the reputation scoring module located within the gateway server module 120 would constitute the claimed “service engine for analyzing a reputation vector for the Web site.” Hubbard-495 further discloses that the URL analysis module 149 “may comprise a reputation scoring module (such as an embodiment of the reputation scoring module 331 of FIG. 12) that generates *a score* or other data representative of a reputation of the URL.” *Id.*, 9:29-33. Hubbard-495 discloses that “the reputation scoring module 331 may generate the scores based on information about the URL of the web page (e.g., scoring criteria).” *Id.*, 18:55-19:22. As I discussed above with respect to element 1[c2], and for the same reasons, Hubbard-495 discloses a “generating a reputation index for the Web site from the reputation vector.” Thus, in my opinion, Hubbard-495 discloses or renders obvious claim 7[e2].

7[e3] the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location,

146. As I discussed above with respect to element 1[c2], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

7[e4] wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block,

147. As I discussed above with respect to element 1[c3], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

7[e5] wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and

148. As I discussed above with respect to element 1[c4], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

7[e6] wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviors, script block count, picture count, immediate redirect and response latency;”

149. As I discussed above with respect to element 1[c5], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

7[f]: “wherein access to the Internet resource by any of the plurality of client-side devices is determined based on the reputation index exceeding a threshold value established for the local area network.”

150. Hubbard-495, alone or in combination with Hubbard-174 discloses or renders obvious this limitation. This limitation is further obvious as explained in 1[d].

151. As I discussed above with respect to element 7[d], and for the same reasons, Hubbard-495 discloses or renders obvious a “plurality of client-side devices.”

152. Hubbard-495 further describes that “the gateway server module 120 determines whether to allow access to the URL *based at least partly on the reputation score.*” Ex.1004, 27:51-53. Additionally, “the reputation scoring module 331 may provide the reputation score to the content scoring module 330, which may determine a category of the requested URL at least partly based on the reputation score.” *Id.*, 27:54-57. “The gateway server module 120 may rely upon the determined category in combination with the score associated with the URL *when determining whether to allow the user access to the website associated with the requested URL.*” *Id.*, 27:58-61. “Each reputation score may then be unique to each website URL and relied upon by the gateway server module 120 to determine under what circumstances, if any, the user will be allowed to access the content associated with each scored website URL gathered by such types of web feed reader applications.” *Id.*, 10:64-11:2. As shown, Hubbard-495 contemplates determining whether access is allowed to the requested URL based on the reputation score.

153. Hubbard-495 also discloses that the policy module 142 “can be configured to implement different policies depending on the score, the user, or any other applicable parameters.” Ex.1004, 11:27-29. The policy module 142 may be

“used to ***implement network policies*** regarding how certain content will be handled by the gateway server module 120 or by a firewall or some other security software installed within the network 110.” *Id.*, 6:54-58. It would have been obvious to a POSITA that the gateway server module 120 and the policy module 142 of Hubbard-495, being part of the same local area network as the client workstations 116, would have a “network policy” that would apply within the network 110, which as discussed above with respect to element 1[a] may be a local area network.

154. Hubbard-495 describes that the URL analysis module 149 may generate a “***sufficiently high reputation score*** (e.g., indicative of a good reputation or low likelihood of active or other targeted content) for an uncategorized URL ***so that the policy module 142 allows access to the uncategorized URL*** when the policy module 142 would otherwise not allow access to the URL.” Ex.1004, 11:29-36. A POSITA would have understood that a “sufficiently high” reputation score would require some metric for which a reputation score would be “sufficient,” such as a threshold value in which reputation values over this threshold would be sufficient to allow access. Hubbard-495 also contemplates the use of thresholds in the scoring process, where, e.g., “categories corresponding to each score above a given ***threshold*** are identified with the URL.” *Id.*, 21:1-4. *See also, e.g., id.*, 19:45-51, 19:55-61, 20:65-21:1 (“In one embodiment, ***URLs with scores above a threshold*** are identified with a category, e.g., malicious, regardless of, or in addition to, the

category identified by content analysis.”), 21:1-10. In my opinion, this would qualify as the claimed “wherein access to the Internet resource by any of the plurality of client-side devices is determined based on the reputation index exceeding a threshold value established for the local area network.” Thus, in my opinion, Hubbard-495 discloses or renders obvious claim 7[f].

H. Claim 8

“The system according to claim 7, wherein the client-side device is a personal computer or a PDA.”

155. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

156. Hubbard-495 discloses that the workstations 116 or the client computers 116 “may be desktop computers, **notebook computers, handheld computers** or the like” that may “be loaded with operating systems that allow them to utilize the network through various software modules such as web browsers, e-mail programs, or the like.” Ex.1004, 4:44-49. A POSITA would have understood that “desktop computers, notebook computers, handheld computers or the like” would constitute “personal computers” or “PDAs.” The ‘386 Patent does not define “PDA,” but a POSITA would understand that “PDA” stands for “Personal Digital Assistant,” which is a handheld computer with e-mail capabilities; hence, a POSITA would understand “notebook computers” in Hubbard-495 to encompass “personal computers,” and “handheld computers” in Hubbard-495 to encompass PDAs. It

would also have been obvious to a POSITA that when discussing “desktop computers, notebook computers, handheld computers *or the like*,” PDAs could also be client devices that could utilize “web browsers, e-mail programs, or the like” and be connected to a network 110.

I. Claim 9

“The system according to claim 7, further comprising a data collection site accessible over the Internet, the security site having a crawler for collecting a plurality of factors from a plurality of Internet resources.”

157. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

158. As I discussed above with respect to claim 5, the database management module 114 of Hubbard-495 qualifies as the claimed “data collection site.” The ‘386 Patent describes a reputation generating site 70 or data collection site 70, “preferably having a crawler 71 and a database 72”, that “provides reputation indices to service engine 52 of the security appliance 50.” Ex.1001, 10:42-58. In my opinion, a POSITA would have understood that “the security site having a crawler for collecting a plurality of factors from a plurality of Internet resources” refers to the “data collection site” which also preferably has a crawler.

159. As seen below in Fig. 7, Hubbard-495 discloses the database management module 114 may also include a URL/content database 180 and a collection system 182.

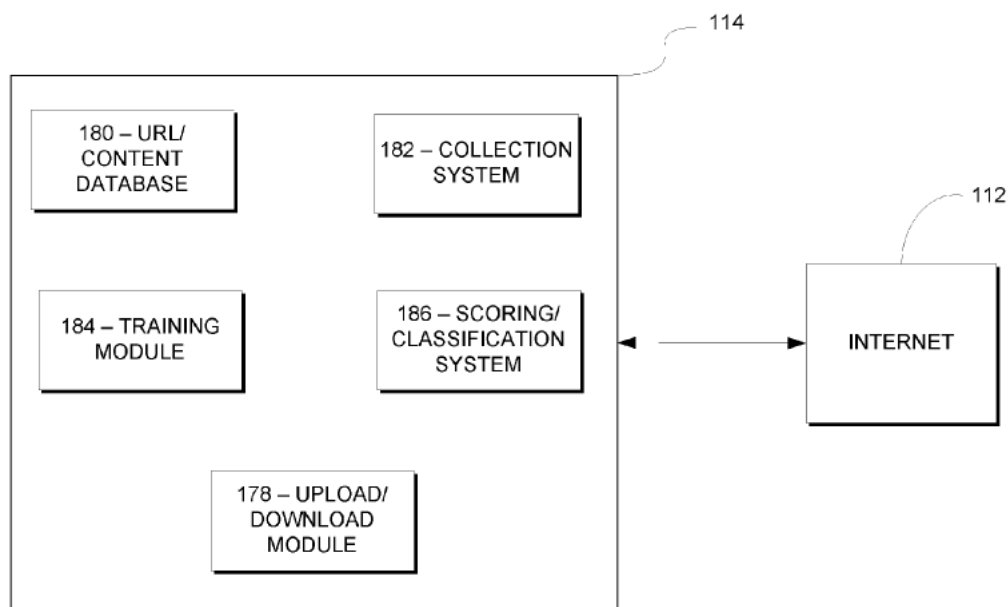


FIG. 7

Ex.1004, Fig. 7.

160. The collection system 182 “may include a database management collection module 190” which may include various subsystems such as the webcrawler module 206 and the honey client module 208. Ex.1004, 12:55-58; 13:59-14:17; *see also id.*, Fig. 9 (shown below). A POSITA would have understood this to constitute the claimed “crawler.”

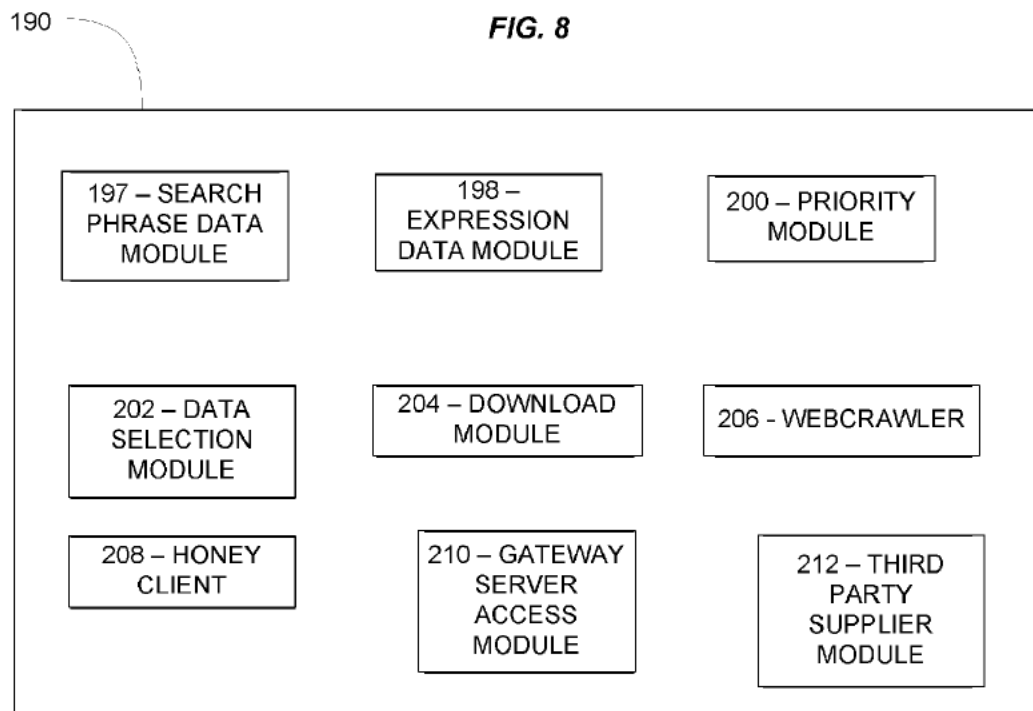


FIG. 9

Ex.1004, Fig. 9.

161. The webcrawler module 206 “may be configured with several concurrent processes that allow the module to simultaneously *crawl many websites* and report the visited URLs back to the database management module URL/content database 180.” Ex.1004, 14:3-8. The database management collection module 190 may also include a honey client module 208 “configured to mimic the behavior of a web browser” that may “visit the web sites and track the behavior of the websites and *download the content back to the URL/content database 180 for further analysis.*” *Id.*, 14:9-18. Hubbard-495 discloses both passive honey miners and active honey miners, with passive honey miners “similar to a web crawler,” but

“unlike the web crawler above which merely visits the website and reports the URL links available from that site, the passive honey miners may be configured to *download the page content and return it to the control servers 220 for insertion into the URL database 180.*” *Id.*, 14:58-64. A POSITA would have understood that the downloading of content from the websites would constitute collecting that content. The collected data may also include “a URL 232”, “the page content 234 for a particular URL,” as well as an “indication of whether the URL has executable content 236 within its page code.” *Id.*, 16:1-20. A POSITA would have understood that the collected data would encompass information about the URL (e.g., scoring criteria) of the web pages visited, e.g., a plurality of factors from a plurality of Internet resources.

162. In my opinion, Hubbard-495 discloses or renders obvious “a data collection site accessible over the Internet, the security site having a crawler for collecting a plurality of factors from a plurality of Internet resources.”

J. Claim 10

“The system according to claim 9, wherein the data collection site has a database containing reputation indices for a plurality of Internet resources.”

163. In my opinion, Hubbard-495 discloses or renders obvious this limitation. As I discussed above with respect to claim 5, the database management module 114 of Hubbard-495 qualifies as the claimed “data collection site.” As seen

in FIG. 7, “the database management module 114 may also include a database management module URL/content database 180.” Ex.1004, 12:25-26.

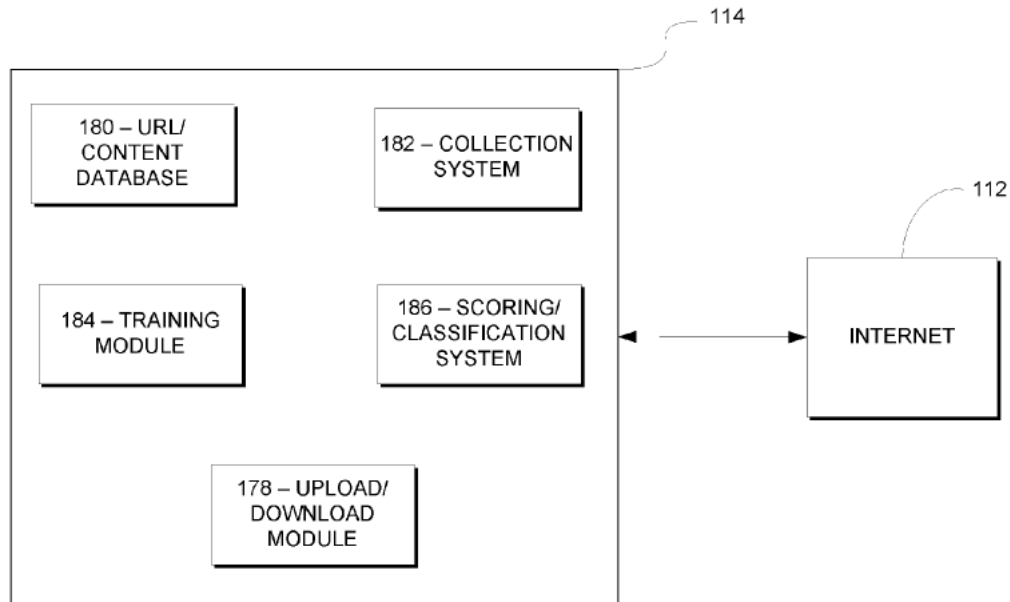


FIG. 7

Ex.1004, Fig. 7.

164. The database management module URL/content database 180 “may take the form of a data warehouse which stores URL strings and information about *URLs* that have been collected by a collection system 182.” Ex.1004, 12:27-30. A POSITA would have understood that the database contains information regarding a plurality of Internet resources, as Hubbard-495 discloses a database containing information about multiple URLs, which a POSITA would have known are Internet resources.

165. Hubbard-495 also teaches that the “database management module URL/content database 180 may be a relational database that is indexed to provide quick and effective searches for data.” Ex.1004, 12:31-33. Further, Hubbard-495 discloses that “[a]dvantageously, a remotely located reputation scoring module may be shared by multiple gateway server modules 120.” *Id.*, 9:61-63; *see id.* at Figs. 7, 12. As I discussed above with respect to claim 6, it would have been obvious to a POSITA to store reputation scores in the database to “provide quick and effective searches for data” and to minimize repeated calculations if a reputation score is requested for the same URL. It would further have been obvious to a POSITA to store reputation scores in the database management module URL/content database 180, along with other information about the URLs, for efficiency. Storing calculation results—such as at the database management module of Hubbard-495—was a common technique in software design at the time.

166. Further, it would have been obvious to a POSITA that it would be advantageous to store the reputation scores in the database already containing information about multiple URLs. Indeed, Hubbard-495 contemplates storage of information about URLs which is used to determine whether access is allowed or denied, disclosing that “the static content classification module 328 **stores** the categories of each URL to a categorized web page database. In one embodiment, the URL database 180 includes the categorized web page database.” Ex.1004, 27:5-9.

Hubbard-495 further discloses that “[o]nce *the data has been collected and stored*, the process moves to block 1410 where the URL request is either allowed or blocked based on the policies indicated in the policy module 142.” *Id.*, 23:60-64.

K. Independent Claim 11

167. In my opinion, Hubbard-495, alone or in view of Hubbard-174, discloses or renders obvious claim 11.

11[preamble]: “A method for controlling access to an Internet resource, the method comprising:”

168. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[preamble].

11[a]: “transmitting a request for an Internet resource from a Web browser for a client-side device of a local area network, the Internet resource residing at a first server;”

169. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[a]. A POSITA would have understood that an “Internet-enabled client application” would necessarily encompass a “Web browser,” which is a type of “Internet-enabled client application.” Hubbard-495 also discloses that the client computers 116 may “be loaded with operating systems that allow them to utilize the network through various software modules such as *web browsers*, e-mail programs, or the like.” Ex.1004, 4:46-49.

11[b]: “receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet;”

170. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[b].

11[c]: “accessing a reputation index for the Internet resource, the reputation index based on a reputation vector for the Internet resource comprising security history, legitimacy, behavior, associations and location, wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block, wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviours, script block count, picture count, immediate redirect and response latency;”

171. In my opinion, Hubbard-495 discloses or renders obvious this element, as I explain below.

11[c1]: “accessing a reputation index for the Internet resource,”

172. Hubbard-495 discloses that “the reputation scoring module 331 may generate the scores based on information about the URL of the web page (e.g., scoring criteria).” Ex.1004, 18:61-19:22. The gateway server module 120 “determines whether to allow access to the URL based at least partly on the reputation score.” *Id.*, 27:51-53. “Each reputation score may then be unique to each

website URL and relied upon by the gateway server module 120 to determine under what circumstances, if any, the user will be allowed to access the content associated with each scored website URL.” *Id.*, 10:64-11:2. A POSITA would have understood that the gateway server module 120 would need to access the generated reputation score in order to rely upon the score to determine whether access would be allowed for the requested web content, which a POSITA would have understood to be an Internet resource. Thus, in my opinion, Hubbard-495 discloses or renders obvious claim 11[c1].

11[c2]: “the reputation index based on a reputation vector for the Internet resource comprising security history, legitimacy, behavior, associations and location,”

173. As I discussed above with respect to element 1[c2], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

11[c3]: “wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block,”

174. As I discussed above with respect to element 1[c3], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

11[c4]: “wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of: age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and”

175. As I discussed above with respect to element 1[c4], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

11[c5]: “and wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviours, script block count, picture count, immediate redirect and response latency;”

176. As I discussed above with respect to element 1[c5], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

11[d]: “determining if the reputation index for the Internet resource is at or above a threshold value established for the local area network; and”

177. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to element 1[c1].

11[e]: “transmitting a decision transmission to the Web browser of the client-side device, the decision transmission allowing or denying access to the Internet resource.”

178. In my opinion, Hubbard-495, either alone or in combination with Hubbard-174, discloses or renders obvious this element for the same reasons that I explained with respect to element 1[d]. A POSITA would have understood that an “Internet-enabled client application” would necessarily encompass a “Web browser,” which is simply a type of “Internet-enabled client application.” Hubbard-495 also discloses that the client computers 116 may “be loaded with operating systems that allow them to utilize the network through various software modules such as *web browsers*, e-mail programs, or the like.” Ex.1004, 4:46-49.

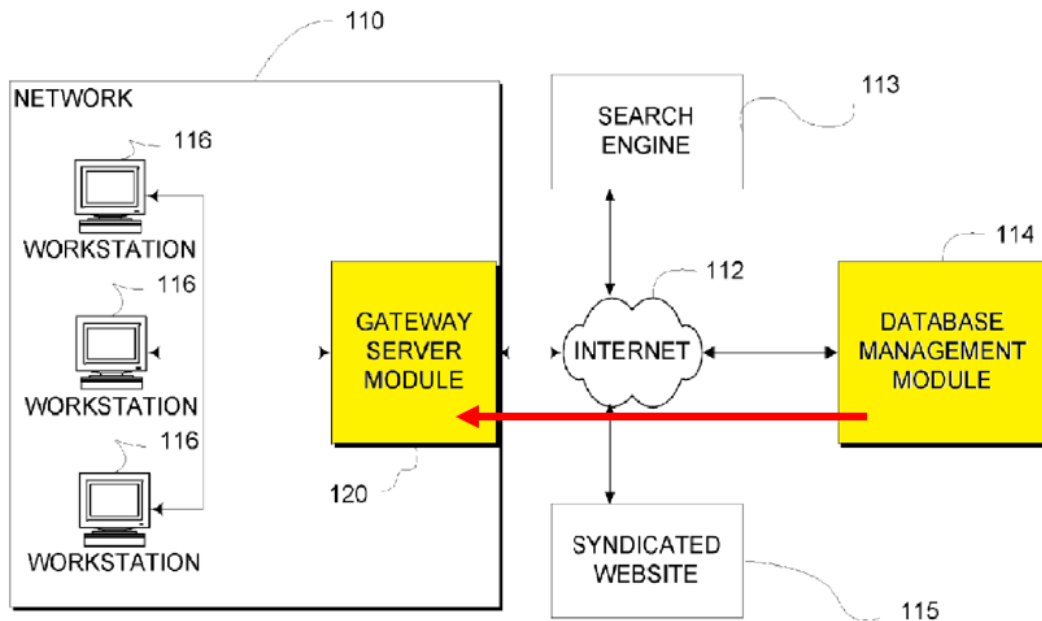
L. Claim 12

“The method according to claim 11, further comprising transmitting the reputation index over the Internet to the local area network.”

179. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

180. Hubbard-495 discloses that the URL analysis module 149 “may rely upon a remotely located reputation scoring module,” which may be a reputation scoring module 331 “located at the database management module 114.” Ex.1004, 9:43-49. The database management module 114 “may be located remotely *(accessible via Internet 112) from the network 110* and its associated workstations 116.” *Id.*, 12:5-8. The URL analysis module 149 “communicates the requested URL to the database management module 114 for scoring by the remotely located reputation scoring module 331” and the database management module 114 “*communicates the results of the analysis performed by the reputation scoring module 331 back to the gateway server module 120.*” *Id.*, 9:49-54. A POSITA would have understood that the “results of the analysis performed by the reputation scoring module 331” would have been the results of “scoring by the remotely located reputation scoring module 331.” Thus, Hubbard-495 discloses that the remotely located reputation scoring module 331 of the database management module 114 generates reputation scores and sends them back to the gateway server module. As

the database management module 114 is connected to the gateway server module 120 and the local area network 110 via the Internet, a POSITA would have understood that this involves transmitting the “results of the analysis performed by the reputation scoring module” in the database management module over the Internet to the local area network 110.



Ex.1004, Fig. 1. (annotated).

M. Independent Claim 13

181. In my opinion, Hubbard-495, alone or in view of Hubbard-174, discloses or renders obvious claim 13, as I explain below.

13[preamble]: “A method for building a reputation database for Internet resources, the method comprising:”

182. In my opinion, Hubbard-495 discloses or renders obvious the preamble.

183. Hubbard-495 discloses that “the database management module 114 may also include a database management module URL/content database 180.” Ex.1004, 12:25-26. Hubbard-495 also teaches that the “database management module URL/content database 180 may be a relational database that is indexed to provide quick and effective searches for data.” *Id.*, 12:31-33. As I discussed above with respect to Claim 6, and for the same reasons, it would have been obvious to a POSITA to store a reputation score in the database to “provide quick and effective searches for data” and to minimize repeated calculations if a reputation score is requested for the same URL. In my opinion, Hubbard-495’s database management module URL/content database 180 that stores reputation scores for URLs would qualify as the claimed “reputation database for Internet resources.”

184. The database management module URL/content database 180 “may take the form of a data warehouse which stores URL strings and information about URLs that have been collected by a collection system 182 which will be described in further detail below.” Ex.1004, 12:27-31. Hubbard-495 further discusses that the “database management module URL/content database 180 may be a relational database that is indexed to provide quick and effective searches for data.” *Id.*, 12:31-33. As I discussed above with respect to claim 9, the collection system 182 may include a webcrawler module 206 and a honey client module 208 that adds information to the database management module URL/content database 180. The

URL/content database 180 “may include data such as URL strings, the content associated with those strings, information about how the content was gathered (e.g., by a honey client, by a customer submission, etc.), and possibly the date in which the URL was written into the URL/content database 180.” *Id.*, 12:36-42. A POSITA would have understood that since Hubbard-495 describes methods to populate such databases, and possible configurations, a method to build a reputation database is described.

13[a]: collecting a plurality of factors for an Internet resource site to populate a reputation vector for the Internet resource to perform reputation analysis of the Internet resource;”

185. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

186. Hubbard-495 discloses the collection system 182, which “may include a database management collection module 190 which is coupled (either directly or indirectly) to a data mining module 192.” Ex.1004, 12:55-58. The database management collection module 190 “may be used by the database management module 114 to *collect data for the URL/content database 180 about URLs* that have not been categorized.” *Id.*, 12:58-61. The database management collection module 190 may “be associated with one or more collection sources 194 from which it may *collect data about URLs.*” *Id.*, 12:64-66. Collection sources 194 may take various forms, such as “active and passive honeypots and honey clients, data analysis of

logging databases 144 stored on the gateway server module 120 to identify applications, URLs and protocols for collection,” “webcrawling applications that search the Internet 112 for particular keywords or search phrases within page content,” and “URLs and IP addresses data mined from a DNS database to identify domains that are associated with known malicious IP addresses.” *Id.*, 12:66-13:10. In my opinion, Hubbard-495’s collection system 182 performs the claimed “collecting a plurality of factors for an Internet resource site.”

187. As I discussed above with respect to element 1[c], Hubbard-495’s “***information about the URL of the web page (e.g., scoring criteria)***” including “one or more of URL content strings,” “meta tag categories associated with other URLs associated with the URL host (e.g., such as in the URL database 180 or the processed web page properties database 324),” and “categories associated with other URLs of the URL host (e.g., in the URL database 180)” would qualify as the claimed “reputation vector comprising a plurality of factors for the Internet resource.” Ex.1004, 18:53-19:22. The database management module URL/content database 180 “may take the form of a data warehouse which stores ***URL strings and information about URLs*** that have been ***collected by a collection system 182.***” *Id.*, 12:27-31. A POSITA would have understood that that the information collected by the collection system 182 and stored in the URL database 180 would be used to populate the information about the URL of the web page (e.g., scoring criteria) and

constitute the claimed “reputation vector for the Internet resource” comprising a *plurality of factors* for the Internet resource.

188. Hubbard-495 discloses that the reputation scoring module 331 “may *generate the scores based on information about the URL of the web page* (e.g., scoring criteria)” such as “URL content strings,” “meta tag categories associated with other URLs associated with the URL host,” and “categories associated with other URLs of the URL host.” Ex.1004, 18:61-19:22. A POSITA would have further understood that generating a reputation score would inherently constitute performing a reputation analysis of the Internet resource, and that that the reputation scoring module 331 uses information about the URL of the web page (e.g., scoring criteria) which encompasses information collected by the collection system 182 and stored in the URL database 180 to generate a reputation score. For these reasons and as described further in 13[d], Hubbard-495 discloses “perform[ing] reputation analysis of the Internet resource.” In my opinion, this would qualify as the claimed “collecting a plurality of factors for an Internet resource site to populate a reputation vector for the Internet resource to perform reputation analysis of the Internet resource.”

13[b]: “receiving the plurality of factors for the Internet resource at a data collection site;”

189. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

190. As I discussed above with respect to claim 5, the database management module 114 of Hubbard-495 qualifies as the claimed “data collection site.” Hubbard-495’s database management module 114 may “also include a database management module URL/content database 180”, that “may take the form of a data warehouse which *stores URL strings and information about URLs that have been collected by a collection system* 182.” Ex.1004, 12:25-31. As I described above for element 13[a], the information collected by the collection system 182 would constitute the claimed “plurality of factors for the Internet resource.” A POSITA would have understood that storing the information that has been collected by a collection system inherently requires receiving the information. Because the database management module 180 of the database management module 114 stores, and thus receives, the information about URLs that have been collected by a collection system, Hubbard-495 discloses “receiving the plurality of factors for the Internet resource at a data collection site.”

13[c]: “constructing a reputation vector for the Internet resource at the data collection site, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location, wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block, wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of age of: a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and wherein the behavior factor

of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviours, script block count, picture count, immediate redirect and response latency;”

191. In my opinion, Hubbard-495 discloses or renders obvious this limitation, as I explain below.

13[c1]: “constructing a reputation vector for the Internet resource at the data collection site,”

192. In my opinion, Hubbard-495 discloses or renders obvious “constructing a reputation vector for the Internet resource at the data collection site.” As I discussed above with respect to elements 13[a] and 13[b], the database management module 114 collects and stores “information about the URL of the web page (e.g., scoring criteria)” or the claimed “plurality of factors for the Internet resource.” Hubbard-495 teaches that the “database management module URL/content database 180 may be a relational database that is indexed to provide quick and effective searches for data.” Ex.1004, 12:31-33. A POSITA would have understood that indexing the “plurality of factors for the Internet resource” *by* URLs would be a logical and efficient way to store the information, and to compile together the various collected information for a discrete URL for ease and efficiency. A POSITA would further have understood that the database would store information about the URL (plurality of factors) as a vector (e.g., database record) used for determining the reputation of the Internet resource. In my opinion, this would qualify as the

“constructing a reputation vector for the Internet resource at the data collection site that comprises a plurality of factors for the Internet resource.”

13[c2]: “the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location,”

193. As I discussed above with respect to element 1[c2], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

13[c3]: “wherein the location factor of the plurality of factors for the reputation vector comprises two or more of: country of domain registration, country of service hosting and country of an internet protocol address block,”

194. As I discussed above with respect to element 1[c3], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

13[c4]: “wherein the legitimacy factor of the plurality of factors for the reputation vector comprises two or more of age of: a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and”

195. As I discussed above with respect to element 1[c4], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

13[c5]: “wherein the behavior factor of the plurality of factors for the reputation vector comprises at least one of: plurality of run-time behaviours, script block count, picture count, immediate redirect and response latency;”

196. As I discussed above with respect to element 1[c5], and for the same reasons, Hubbard-495 discloses or renders obvious this limitation.

13[d]: “analyzing the reputation vector to generate a reputation index for the Internet resource based on the analysis of the plurality of factors and the reputation classifier;”

197. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

198. Hubbard-495 discloses that “the reputation scoring module 331 may generate the scores based on information about the URL of the web page (e.g., scoring criteria).” Ex.1004, 18:61-19:22. In my opinion, Hubbard-495 discloses or renders obvious “analyzing the reputation vector to generate a reputation index for the Internet resource based on the analysis of the plurality of factors” for the same reasons as discussed above with respect to element 1[c].

199. Hubbard-495 discloses that “certain combinations of the scoring criteria are particularly useful for identifying active or other target content” because the set of criteria may cover “a number of different ways that targeted active content might be created to avoid detection.” Ex.1004, 19:62-66. “[M]ultiple scores associated with different categories” may be “assigned to each URL, and the categories corresponding to each score above a given threshold are identified with the URL.” *Id.*, 21:1-4. For example, the static content classification module 328 may “compare[] the properties for one or more web pages with definitions from the definitions database 326” and “when a web page matches a particular definition, the web page is identified with one or more categories associated with the definition.”

Id., 18:3-9. Hubbard-495 discloses that “categories can refer to any type of classification.” *Id.*, 18:38-39.

Definitions Database

FIG. 13C

Expression	Category
Content-Type = HTML and “<javascript>” > 2	Malicious
<object> = 6 and URL matches “a.com”	Gambling
“data=,*\txt” = “data=xyx333.txt	Spyware

Ex.1004, Fig. 13C.

200. As seen in FIG. 13C, the categories may include such categories as “malicious,” “gambling,” or “spyware.” In my opinion, a POSITA would have understood these categories to provide a classification related to the likelihood of the URL or Internet resource being malicious or spyware, i.e., a reputation classifier.

201. It would have been obvious to a POSITA to use the category classifications and the information about the URL of the web page (e.g., scoring criteria) of Hubbard-495 to generate a reputation score, as combining the two would cover a number of different ways that targeted active content might be created to avoid detection.

13[e]: “storing the reputation index for the Internet resource at the data collection site; and”

202. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to claim 6.

13[f]: “transmitting the stored reputation index to a local area network upon request for managing access to the Internet resource.”

203. In my opinion, Hubbard-495 discloses or renders obvious this element. In my opinion, Hubbard-495 discloses or renders obvious “transmitting the reputation index to a local area network” for the same reasons that I explained with respect to claim 12.

204. Hubbard-495 discloses a “network request” that includes “*user requests* for content via the search engine 113 or requests sent via a web feed reader to the syndicated website 115.” Ex.1004, 5:51-54. “[*R*]requests originating within the workstations 116 of the network 110 may be required to pass through the gateway server module 120 as they proceed to the Internet.” *Id.*, 6:24-26. Hubbard-495 discloses that the URL analysis module 149 “*communicates the requested URL to the database management module 114 for scoring by the remotely located reputation scoring module 331*” and that the database management module 114 “*communicates the results of the analysis performed by the reputation scoring module 331 back to the gateway server module 120.*” *Id.*, 9:49-54. “[T]he gateway server module 120 *determines whether to allow access to the URL based at least partly on the reputation score.*” *Id.*, 27:51-53. It would have been obvious to a POSITA that the request for a reputation score so that the gateway server module

120 could determine whether to allow access would qualify as a request for managing access to the Internet resource.

205. As I discussed above for claim 6 and element 13[e], a POSITA would have understood that the database management module 114 could store the reputation score for a URL, or Internet resource. Element 13[e] discusses “storing the reputation index for the Internet resource at the data collection site.” It would have been obvious to a POSITA to transmit the the reputation score stored in the database if the reputation index was already stored in the database to minimize repeat calculations if a reputation score was requested for the same URL, or Internet resource, as this would be more efficient. This limitation is further obvious as explained in 1[d].

N. Claim 15

“The method according to claim 13, further comprising obtaining the plurality of factors for the Internet resource using a crawler.”

206. In my opinion, Hubbard-495 discloses or renders obvious this element for the same reasons that I explained with respect to claim 9.

O. Claim 16

“The method according to claim 15, wherein obtaining the plurality of factors for the Internet resource comprises accessing the Internet service, analyzing a plurality of HTML documents for the Internet resource, and crawling a plurality of linked Internet resources of the plurality of HTML documents for Internet resource.”

207. In my opinion, Hubbard-495 discloses or renders obvious this limitation.

208. As discussed above with respect to claim 9, the collection system 182 “may include a database management collection module 190” which may include various subsystems such as the webcrawler module 206 and the honey client module 208. Ex.1004, 12:55-58; 13:59-14:17. “The webcrawler module 206 may be configured with several concurrent processes that allow the module to simultaneously *crawl many websites* and *report the visited URLs* back to the database management module URL/content database 180.” *Id.*, 14:3-8. A POSITA would have understood that visiting the websites would constitute the claimed “accessing the Internet service” as it would have been necessary to access the Internet server on which the website is stored in order to visit it.

209. The database management collection module 190 may also include a honey client module 208 “configured to mimic the behavior of a web browser” that “may *visit the web sites* and track the behavior of the websites and *download the content back to the URL/content database 180 for further analysis.*” Ex.1004, 14:9-16. Hubbard-495 discloses both passive honey miners and active honey miners, with passive honey miners “similar to a web crawler,” but “unlike the web crawler above which merely *visits the website and reports the URL links available from that site*, the passive honey miners may be configured to *download the page*

content and return it to the control servers 220 for insertion into the URL database 180.” *Id.*, 14:57-64. Hubbard-495 provides for “systems and method of identifying and categorizing web content, including potentially executable web content and malicious content, that is found at locations identified by Uniform Resource Locators (URLs).” *Id.*, 4:3-7. Potentially executable web content may include, for example, “applets, *executable code embedded in HTML or other hypertext documents (including script languages such as JavaScript or VBScript)*, executable code embedded in other documents, such as Microsoft Word macros, or stylesheets,” “documents that execute code in another location such as another web page, another computer, or on the web browser computer itself,” and “a *HTML web page that includes an “OBJECT” element*, and thus can cause execution of ActiveX or other executable components.” *Id.*, 4:10-22. Malicious web content may refer to “interactive content such as ‘phishing’ schemes in which a HTML form or other web content is designed to appear to be provided by another, typically trusted, web site such as a bank, in order to deceive the user into providing credentials or other sensitive information to an unauthorized party.” *Id.*, 4:28-35. A POSITA would have understood that such reporting of the websites and the web content on those websites back to the URL/content database 180 for further analysis would qualify as the claimed “analyzing a plurality of HTML documents for the Internet

resource” as it was well known at the time that web pages were documents, commonly written in HTML, that are viewed in an Internet browser.

210. “In some embodiments, the miners 222 *may be configured to follow links on the URL content* so that in addition to visiting URLs specified by the control server 220, the miners may visit content that it linked to those URLs.” Ex.1004, 15:9-13. Hubbard-495 further discloses configuring miners to follow links on the URL content. A POSITA would have understood such following of links on the URL content would constitute “crawling a plurality of linked Internet resources of the plurality of HTML documents for Internet resource.”

P. Claim 18

“The method according to claim 13, wherein the stored reputation index is transmitted over the Internet to a local area network.”

211. In my opinion, Hubbard-495 discloses or renders obvious “transmitting the stored reputation index to a local area network” for the same reasons that I explained with respect to elements 13[e] and 13[f]. As discussed in those sections, a POSITA would have understood that the reputation scores could be stored in the database management module URL/content database 180 of the database management module 114, along with other information about the URLs, for efficiency, and that data backups and mirroring were common techniques in software design at the time.

212. As I discussed above with respect to claim 5, the database management module 114 of Hubbard-495 qualifies as the claimed “data collection site.” Hubbard-495 discloses that the database management module 114 “may be located remotely (*accessible via Internet 112*) *from the network 110* and its associated workstations 116.” Ex.1004, 12:5-8. As the database management module 114 is connected to the gateway server module 120 and the local area network 110 via the Internet, a POSITA would have understood that information sent from the database management module 114 to the gateway server module 120 would be over the Internet to the local area network.

IX. ANALYSIS OF HUBBARD-495 IN COMBINATION WITH HUBBARD-174 AND ALPEROVITCH FOR CLAIMS 1-13, 15, 16, AND 18

213. In my opinion, Hubbard-495 in view of Hubbard-174 and Alperovitch discloses or renders obvious claims 1–13, 15, 16, and 18 of the ‘386 Patent for the reasons set forth below.

A. Independent Claims 1, 7, 11, and 13

214. Hubbard-495 in combination with Hubbard-174 discloses or renders obvious each limitation of independent claims 1, 7, 11, and 13, as I explain above in Sections IX.A, G, K, M.

215. To the extent that Hubbard-495 in combination with Hubbard-174 does not expressly disclose “the reputation vector [comprising a plurality of factors] for the Internet resource comprising security history”, Alperovitch discloses it or

renders it obvious. Alperovitch discloses using behavioral attributes to derive a “behavioral profile” that includes “behavioral events relevant to defining the disposition of an entity matching the profile” which “include, ‘Spammer’, ‘Scammer’, and ‘Legitimate Sender’”. Ex.1006, [0048]. “Events and/or evidentiary attributes relevant to each profile define appropriate entities to which a profile should be assigned.” *Id.* This may include “specific set of sending patterns, blacklist events, or specific attributes of the evidentiary data.” *Id.* A POSITA would have understood that blacklist events would be previous security events that would flag a site for placement on the “blacklist” or a list of entities that are blocked from communicating with or logging into a computer, site, or network, and that “a specific set of sending patterns” would reference historical data about the entity’s prior communications. A POSITA would have understood that Alperovitch discloses historical events that flag security risks, which in my opinion would qualify as the claimed “security history” factor of the “plurality of factors for the Internet resource.”

Motivation to Combine Hubbard-495, Hubbard-174 and Alperovitch

216. In my opinion, a POSITA would have been motivated to combine Hubbard-495, Hubbard 174 and Alperovitch because a POSITA would have recognized them as analogous art to the ‘386 Patent at least because each is from the same field of endeavor, namely methods and systems, to determine conditions under

which access to an Internet resource is allowed. *See, e.g.*, Ex.1004, 2:33-61; Ex.1005, 1:7-9; Ex.1006, [0073].

217. Hubbard-495 discloses that the reputation scores may be generated “based on information about the URL of the web page (e.g., scoring criteria) such as... the name server that provides the IP address of the URL host or its domain, the network or blocks of IP addresses associated with the host IP address or its domain, active content previously identified with reference to the host, the hosts’ domain, or the URL path, historical data associated with the IP address of the URL host, creation date or expiration date associated with the registration of the URL host,” among others. Ex.1004, 18:53-19:22. Likewise, Hubbard-174 discloses a “scoring/classification system 186 which utilizes the definitions and properties created by the training system 184 to provide a score or classification (e.g., a categorization) to web content so that the categorization may be delivered via the upload/download module 178 to gateway server modules 120.” Ex.1005, 10:3-9. A POSITA would have understood that these listed examples would be non-exclusive examples of information relevant to the reputation of the URL. Alperovitch discloses “methods and systems for assigning reputation to communications entities” as “hackers often attempt to use various network-type applications (e.g., e-mail, web, instant messaging (IM) etc.) to create a pre-textual connection with the networks 110 a-n in order to exploit security holes created by these various

applications using entities 130 a-e.” Ex.1006, Abstract; [0032]. Hubbard-495 and Hubbard-174 recognize this risk in “phishing” messages in the context of malicious web content, “in which a HTML form or other web content is designed to appear to be provided by another, typically trusted, web site such as a bank, in order to deceive the user into providing credentials or other sensitive information to an unauthorized party.” Ex.1004, 4:29-34; Ex.1005, 4:10-15. In my opinion, a POSITA at the relevant time would have understood that the systems of Hubbard-495 and Hubbard-174 could be improved by combination with Alperovitch’s related reputation analysis system. A POSITA would have had a reasonable expectation of success in combining Hubbard-495 and Hubbard-174, as the combination would have predictably enhanced the system of Hubbard-495 and Hubbard-174 by providing, for example, further “information about the entities 300a-c sending and receiving messages, including, transmission patterns, volume, or whether the entity has a tendency to send certain kinds of message (e.g., legitimate messages, spam, virus, bulk mail, etc.), among many others” on which to base a reputation score. Ex.1006 [0037].

B. Dependent Claims 2, 3, 5, 6, 8, 9, 10, 12, 15, 16, and 18

218. Hubbard-495 in combination with Hubbard-174 alone or in view of Alperovitch, discloses or renders obvious the additional limitations of claims 2, 3, 5, 6, 8, 9, 10, 12, 15, 16, and 18, as I explain above in Sections IX.A. and X.A.

C. Dependent Claim 4

219. Hubbard-495 in combination with Hubbard-174 alone or in view of Alperovitch, discloses or renders obvious the method of claim 1 for the reasons I explained above in Sections IX.A and X.A. To the extent that Hubbard-495 in combination with Hubbard-174 does not disclose “the Internet resource is one of a Web site, an email server and a voice communications server,” Alperovitch discloses or renders it obvious. Alperovitch discloses that reputation engines can assist in identifying “malicious communications,” examples of which include “malicious communications in an email, instant messaging, VoIP, SMS or other communication protocol system.” Ex.1006 [0034]. Because an IP phone establishes a connection to the local network upon establishing a VoIP call, the “connection can be exploited similarly to the way *e-mail, web, instant messaging, or other internet applications* can be exploited for providing unregulated connect to a network.” *Id.*, [0072]. A POSITA would have understood due to the similarities in which VoIP calls and e-mail, web, instant messaging, or other internet applications interact with the local area network, Hubbard-495’s system in view of Alperovitch would extend similarly to voice communications servers, in addition to Web sites and email servers.

X. SECONDARY CONSIDERATIONS

220. I have not seen any evidence of secondary considerations that would support a determination of non-obviousness of the claimed subject matter in the ‘386

Patent. To the extent the Patent Owners or its expert provides such secondary considerations, I reserve the right to respond.

XI. CONCLUSION

221. It is my opinion that claims 1-13, 15, 16, and 18 of the '386 Patent are unpatentable.

222. I hereby declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct, that all statements made of my own knowledge are true, and that all statements made on information and belief are believed to be true. I understand that willful false statements are punishable by fine or imprisonment or both. *See* 18 U.S.C. § 1001.

Date: March 31, 2023

Respectfully submitted,



Dr. Michael T. Goodrich