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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 January 2007 (04.01.2007)

PCT

(10) International Publication Number
WO 2007/002002 A1

(51) International Patent Classification:
H04L 12/58 (2006.01)

(21) International Application Number:
PCT/US2006/023847

(22) International Filing Date: 20 June 2006 (20.06.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/157,327 20 June 2005 (20.06.2005) US

(71) Applicant *(for all designated States except US)*: **SYMANTEC Corporation** [US/US]; 350 Ellis Street, MS-MVCA, Mountain View, California 94043 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **JENSEN, Sanford** [US/US]; 1425 Bancroft Way, Berkeley, California 94702 (US).

(74) Agent: **MEYERTONS, HOOD, KIVLIN, KOWERT & GOETZEL, P.C.**; KIVLIN, B. Noel, P.O. Box 398, Austin, Texas 78767-0398 (US).

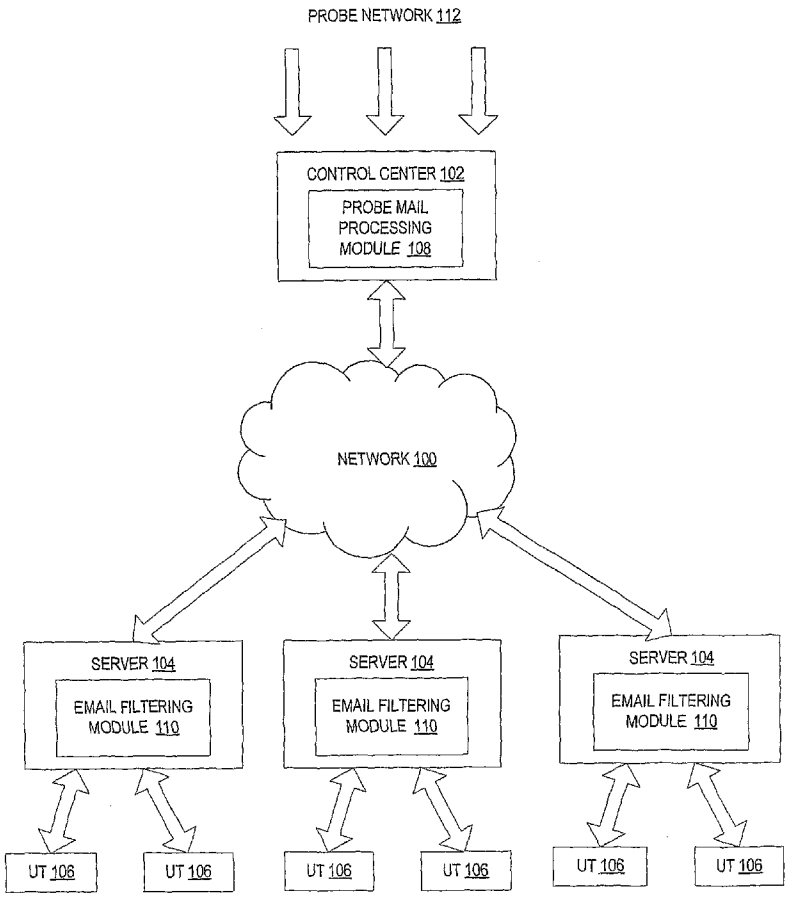
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ,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

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

Published:
— *with international search report*

[Continued on next page]

(54) **Title:** METHOD AND APPARATUS FOR GROUPING SPAM EMAIL MESSAGES



(57) **Abstract:** A method and system for grouping spam email messages are described. In one embodiment, the method includes receiving probe email messages indicative of spam and modifying the probe email messages to reduce noise. The method further includes comparing the probe email messages using fuzzy logic to identify similar email messages, and creating groups of similar email messages. Each of the created groups pertains to a distinct spam attack.

WO 2007/002002 A1

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

METHOD AND APPARATUS FOR GROUPING SPAM EMAIL MESSAGES

FIELD OF THE INVENTION

[0001] The present invention relates to filtering electronic mail (email); more particularly, the present invention relates to creating filters to detect email spam.

BACKGROUND OF THE INVENTION

[0002] In recent years, spam has become a major problem for all Internet users. As the cost of processing power, email address acquisition and email software continue to fall, spam becomes increasingly cost-effective for spammers. Given the negligible cost involved in sending millions of unsolicited email messages, spammers need only capture a small response rate to make a profit. The growth trend of spam shows no sign of abating. According to recent statistics, spam currently accounts for over half of all email traffic in the U.S. This increase in both the volume and percentage of spam is not only worsening a resource drain for IT, it is also affecting how end users view email, which has become the primary form of communication in the enterprise.

[0003] Presently, there are products for filtering out unwanted email messages. However, these products typically fail to effectively compensate for the escalating volumes of spam.

SUMMARY OF THE INVENTION

[0004] A method and system for grouping spam email messages are described. According to one aspect, the method includes receiving probe email messages indicative of spam and modifying the probe email messages to reduce noise. The method further includes comparing the probe email messages using fuzzy logic to identify similar email messages, and creating groups of similar email messages. Each of the created groups pertains to a distinct spam attack.

[0005] Other features of the present invention will be apparent from the accompanying drawings and from the detailed description that follows.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of various embodiments of the invention, which, however, should not be taken to limit the invention to the specific embodiments, but are for explanation and understanding only.

[0007] **Figure 1** is a block diagram of one embodiment of a system for controlling delivery of spam electronic mail.

[0008] **Figure 2** is a block diagram of one embodiment of a probe mail processing module.

[0009] **Figure 3** is a block diagram of one embodiment of a grouping module.

[0010] **Figure 4** is a flow diagram of one embodiment of a process for facilitating detection of spam email messages.

[0011] **Figure 5** is a flow diagram of one embodiment of a process for grouping messages.

[0012] **Figure 6** illustrates a process of grouping messages using fuzzy logic, according to one embodiment of the present invention.

[0013] **Figure 7** is a flow diagram of one embodiment of a process for generating mathematical signatures of documents.

[0014] **Figure 8** is a block diagram of an exemplary computer system.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0015] A method and apparatus for grouping spam email messages are described. In the following description, numerous details are set forth. It will be apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

[0016] Some portions of the detailed descriptions which follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

[0017] It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic)

quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0018] The present invention also relates to apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

[0019] The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

[0020] A machine-readable medium includes any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computer). For example, a machine-readable medium includes read only memory ("ROM"); random access memory ("RAM"); magnetic disk storage media; optical storage media; flash memory devices;

electrical, optical, acoustical or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.); etc.

[0021] **Figure 1** is a block diagram of one embodiment of a system for controlling delivery of spam electronic mail (email). The system includes a control center 102 coupled to a communications network 100 such as a public network (e.g., the Internet, a wireless network, etc.) or a private network (e.g., LAN, Intranet, etc.). The control center 102 communicates with multiple network servers 104 via the network 100. Each server 104 communicates with user terminals 106 using a private or public network.

[0022] The control center 102 is an anti-spam facility that is responsible for analyzing messages indicative of spam, developing filtering rules for detecting spam, and distributing the filtering rules to the servers 104. A message may be indicative of spam because it was collected via a "probe network" 112. In one embodiment, the probe network is formed by fictitious probe email addresses specifically selected to make their way into as many spammer mailing lists as possible. The fictitious probe email addresses may also be selected to appear high up on spammers' lists in order to receive spam mailings early in the mailing process (e.g., using the e-mail address "aardvark@aol.com" ensures relatively high placement on an alphabetical mailing list). The fictitious probe email addresses may include, for example, decoy accounts and expired domains. In addition, a certain percentage of assignable e-mail addresses offered by an ISP or private network may be reserved for use as probe email addresses. The probe network 112 may also receive email identified as spam by users of terminals 106.

[0023] A server 104 may be a mail server that receives and stores messages addressed to users of corresponding user terminals. Alternatively, a server 104 may be a different server (e.g., a gateway of an Internet Service Provider (ISP)) coupled to a mail

server. Servers 104 are responsible for filtering incoming messages based on the filtering rules received from the control center 102. Servers 104 operate as clients receiving services of the control center 102.

[0024] In one embodiment, the control center 102 includes a probe mail processing module 108 that is responsible for identifying spam email messages resulted from distinct spam attacks, generating filters for the distinct spam attacks, and distributing the filters to the servers 104 for detection of spam email resulted from these spam attacks at the customer sites.

[0025] Each server 104 includes an email filtering module 110 that is responsible for storing filters received from the control center 102 and detecting spam email using these filters.

[0026] In an alternative embodiment, each server 104 hosts both the probe mail processing module 108 that generates spam filters and the email filtering module 110 that uses the generated filters to detect spam email.

[0027] **Figure 2** is a block diagram of one embodiment of a probe mail processing module 200. The probe mail processing module 200 includes a probe email parser 202, a noise reduction algorithm 204, a grouping module 206, a filter generator 208 and a filter transmitter 210.

[0028] The probe email parser 202 is responsible for parsing the body of probe email messages.

[0029] The noise reduction algorithm 204 is responsible for detecting data indicative of noise and removing noise from probe email messages. Noise represents data invisible to a recipient that was added to an email message to hide its spam nature. Such data may include, for example, formatting data (e.g., HTML tags), numeric character references,

character entity references, URL data of predefined categories, etc. Numeric character references specify the code position of a character in the document character set. Character entity references use symbolic names so that authors need not remember code positions. For example, the character entity reference "&aring" refers to the lowercase "a" character topped with a ring. Predefined categories of URL data may include, for example, numerical character references contained in the URL and the URL "password" syntax data that adds characters before an "@" in the URL hostname.

[0030] The grouping module 206 is responsible for grouping probe email messages that are likely to result from distinct spam attacks. The grouping module 206 compares probe email messages or their portions (e.g., message headers, message bodies (or portions of message body), message senders, or any combination of the above) to find similar probe email messages. The comparison may be done using regular expressions or mathematical signatures of probe email messages. Mathematical signatures of probe email messages may consist of checksums, hash values or some other data identifying the message content, and may be created using various algorithms that enable the use of similarity measures in comparing different email messages.

[0031] The filter generator 208 is responsible for generating filters for individual groups created by the grouping module 206. A filter may include a mathematical signature of a probe email message, a regular expression characterizing a probe email message, one or more URLs extracted from a probe email message, or any other data characterizing probe email messages resulted from a spam attack.

[0032] The filter transmitter 210 is responsible for distributing filters to participating servers such as servers 104 of **Figure 1**. In one embodiment, each server 104 periodically (e.g., each 5 minutes) initiates a connection (e.g., a secure HTTPS connection)

with the call center 102. Using this pull-based connection, filters are transmitted from the call center 102 to the relevant server 104 for detection of spam email at relevant customer sites.

[0033] **Figure 3** is a block diagram of one embodiment of a grouping module 300. The grouping module 300 includes a feature extractor 302, a signature generator 304, a signature matching algorithm 306, and a group creator 308.

[0034] The feature extractor 302 is responsible for determining the size of probe email messages and expanding a probe email message if its size is below a threshold. In particular, if the feature extractor 302 determines that the size of an email message is below a threshold, it identifies a predefined feature in the probe email message and appends this feature to the probe email message to increase its size. The feature may be appended to the probe email message several times until the size of the probe email message reaches a threshold. The feature may be any data contained in the message that characterizes the content of the message or its sender or recipient. For example, the feature may be a URL, a telephone number, a keyword, a company name or a person's name.

[0035] The signature generator 304 is responsible for generating mathematical signatures for probe email messages. The signature generator 304 may generate mathematical signatures using various algorithms that enable the use of similarity measures in comparing different email messages. For example, the signature generator 304 may generate mathematical signatures using the MD5 algorithm or a character-based algorithm that extracts from a message the most frequently occurring combinations of characters.

[0036] The signature matching algorithm 306 is responsible for comparing signatures of probe email messages to identify similar probe email messages. In one embodiment, the signature matching algorithm 306 operates using fuzzy logic to not only

identify matching probe email messages but also probe email message that are significantly similar and, therefore, are likely to have been originated from the same spam attack.

[0037] The group creator 308 is responsible for grouping similar probe email messages into groups corresponding to distinct spam attacks.

[0038] The grouping module 300 operates in real time, efficiently grouping a large number of probe email messages (e.g., several million messages per day) to identify spam attacks at an early stage and allow for creation of filters that detect spam email messages resulting from these spam attacks at the customer sites.

[0039] **Figure 4** is a flow diagram of one embodiment of a process 400 for facilitating detection of spam email messages. The process may be performed by processing logic that may comprise hardware (e.g., dedicated logic, programmable logic, microcode, etc.), software (such as run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, processing logic resides at a control center 102 of **Figure 1.**

[0040] Referring to **Figure 4**, process 400 begins with processing logic receiving probe email messages indicative of spam (processing block 402). Probe email messages are indicative of spam because they are collected via a probe network.

[0041] At processing block 404, processing logic modifies the probe email messages to reduce noise. In one embodiment, processing logic modifies the probe email messages by removing formatting data, translating numeric character references and character entity references to their ASCII (American Standard Code for Information Interchange) equivalents, and modifying URL data of predefined categories. In one embodiment, formatting data is removed if it does not qualify as an exception. Typically, HTML formatting does not add anything to the information content of a message. However, a few

exceptions exist. These exceptions are the tags that contain useful information for further processing of the message (e.g., tags <BODY>, <A>, <IMG>, and <FONT>). For example, the <FONT> and <BODY> tags are needed for "white on white" text elimination, and the <A> and <IMG> tags typically contain link information that may be used for passing data to other components of the system.

[0042] In one embodiment, processing logic modifies URL data of predefined categories by removing numerical character references contained in the URL, removing additional characters added before an "@" in the URL hostname, and removing the "query" part of the URL, following a string "?" at the end of the URL.

[0043] At processing block 406, processing logic generates mathematical signatures of the modified probe messages. Mathematical signatures of probe email messages may consist of checksums, hash values or some other data identifying the message content, and may be created using various algorithms that enable the use of similarity measures in comparing different email messages. In one embodiment, a mathematical signature for an email message is created using a character-based mechanism that identifies the most frequently used character combinations in a document. This mechanism may use an ASCII character set based on the Roman alphabet or an extended character set that also includes non-ASCII characters to cover international email messages using non-Roman alphabets (e.g., Cyrillic alphabet, Greek alphabet, Arabic alphabet, etc.). One embodiment of a process for creating a mathematical signature of a document will be discussed in more detail below in conjunction with **Figure 7**.

[0044] In one embodiment, if a probe email message is small (e.g., if the size of the message is below a threshold), processing logic expands the size of the email message prior to generating a mathematical signature. The size may be expanded by repeatedly

appending to the email messages an important feature extracted from the email message. Such a feature may be a URL, a telephone number, a keyword or a name contained in the email messages. One embodiment of a process for grouping small message will be discussed in more detail below in conjunction with **Figure 5**.

[0045] At processing block 408, processing logic compares the mathematical signatures of probe email messages using fuzzy logic to identify similar probe email messages. One embodiment of a process for finding similar email messages will be discussed in more detail below in conjunction with **Figure 6**.

[0046] At processing block 410, processing logic creates groups of similar email messages. Each group contains email messages that have likely been originated from a distinct spam attack.

[0047] At processing block 412, processing logic creates a filter for each group of email messages. Each filter is intended to detect spam resulting from a corresponding spam attack at the customer sites. A filter may include a mathematical signature of a probe email message, a regular expression characterizing a probe email message, one or more URLs extracted from a probe email message, or any other data characterizing a probe email message.

[0048] At processing block 414, processing logic distributes the filters to the clients for detection of spam at the customer sites.

[0049] **Figure 5** is a flow diagram of one embodiment of a process 500 for grouping messages. The process may be performed by processing logic that may comprise hardware (e.g., dedicated logic, programmable logic, microcode, etc.), software (such as run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, processing logic resides at a control center 102 of **Figure 1**.

[0050] Referring to **Figure 5**, process 500 begins with processing logic determining whether the size of a new message is below a threshold (processing block 502). A new message may be an email message or any other message or document in electronic form. A threshold may be a minimum message size that enables a meaningful comparison of similarities in messages. If the size of the new message is equal to, or exceeds, the threshold, processing logic proceeds to processing box 510. Otherwise, if the size of the new message is below the threshold, processing logic proceeds to processing box 504.

[0051] At processing box 504, processing logic finds a predefined feature in the new message. In one embodiment, the predefined feature is a URL. Alternatively, the predefined feature may refer to any other data that is known to be an important part of the message content. For example, the predefined feature may be a telephone number, a keyword, a company name, a name of a person, etc.

[0052] At processing block 506, processing logic appends the predefined feature to the message. In one embodiment, in which the predefined feature is a URL, processing logic extracts top-level information from each URL found in the message and adds the extracted data to the end of the message. For example, processing logic may extract from each URL a host name (e.g., www.google.com), ignoring randomized sub-domains, redirections to target URLs, randomized paths, etc.

[0053] At processing block 508, processing logic determines whether the size of the expanded message reaches the threshold. If not, processing logic repeats the addition of the extracted feature to the message (processing block 506) until the message size reaches the threshold. When the message size reaches the threshold, processing logic proceeds to processing block 510.

[0054] At processing block 510, processing logic generates a mathematical signature of the new message. One embodiment of a process for creating a mathematical signature of a message will be discussed in more detail below in conjunction with **Figure 7**.

[0055] Next, processing logic compares the signature of the new message with signatures of other messages from the existing groups (processing block 512) and determines whether the signature of the new message is similar to any other signatures (processing block 514). If so, processing logic groups the signature of the new message with the similar signatures (processing block 518). If not, processing logic creates a new group for this signature that denotes a new spam attack (processing block 516).

[0056] **Figure 6** illustrates a process of grouping messages using fuzzy logic, according to one embodiment of the present invention.

[0057] Referring to **Figure 6**, signature 1 is a mathematical signature of a probe email message initially received by the control center 102. Group 1 contains signature 1 and corresponds to a first spam attack characterized by space 600. Space 600 defines an allowable degree of similarity between signatures of group 1.

[0058] Group 2 contains signature 2 and corresponds to a second spam attack characterized by space 602.

[0059] Group 3 contains signature 3 and corresponds to a third spam attack characterized by space 604. Space 604 overlaps with spaces 600 and 602.

[0060] Subsequently, when signatures of newly received messages are compared with existing signatures, they are grouped depending on their similarities with the existing signatures. If a signature is similar to signatures from multiple groups, it is assigned to the oldest group (i.e., the group that was created earlier than the other group or groups). For

example, signature 6 belongs both to space 600 and space 604. Because group 1 was created before group 3, signature 6 is assigned to group 6.

[0061] **Figure 7** is a flow diagram of one embodiment of a process 700 for generating a mathematical signature of a document. The process may be performed by processing logic that may comprise hardware (e.g., dedicated logic, programmable logic, microcode, etc.), software (such as run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, processing logic resides at a control center 102 of **Figure 1**.

[0062] Referring to **Figure 7**, process 700 begins with processing logic pre-processing a document (processing block 702). In one embodiment, the document is pre-processed by identifying all URLs and email addresses with the document.

[0063] At processing block 704, processing logic divides these URLs and email addresses into elements. Each element is a combination of two or more portions of a relevant URL or email address. For example, an email message may contain the URL "http://www.fake-domain.com/place.html" and the email address "user@badguy.com". The URL may be divided into "http://www.fake-domain.com", "www.fake-domain.com", "fake-domain.com" and "place.html". The email address may be divided into "user@badguy.com" and "badguy.com".

[0064] At processing block 706, processing logic creates a mathematical signature of the document by compiling the elements created at processing block 704 into a list. In one embodiment, the elements are listed in the descending or ascending order. In the example above, the mathematical signature may include a list of elements resulting from the URL and the email address.

[0065] Subsequently, the mathematical signature of the document may be compared with other signatures. In one embodiment, the signature is considered sufficiently similar with some other signature if the two signatures include a certain number or percentage of the same list elements. For example, a signature created for a message containing the URL "http://www.fake-domain.com/otherplace.html" and the email address "otheruser@badguy.com" may be considered sufficiently similar with the signature of the referenced-above message that contains the URL "http://www.fake-domain.com/place.html" and the email address "user@badguy.com" because the two signatures include 4 common list elements (about 67 percent).

[0066] It should be noted that a signature can be created for the message body only, the message header only, a combination of the message body and header, a combination of the message body and some other data (e.g., recipient or sender information), or a combination of the message header and some other data (e.g., recipient or sender information).

[0067] It will be understood by one of ordinary skill in the art that various techniques other than those described above may be used by embodiments of the present invention to generate mathematical signatures of email message.

An Exemplary Computer System

[0068] **Figure 8** is a block diagram of an exemplary computer system 800 that may be used to perform one or more of the operations described herein. In alternative embodiments, the machine may comprise a network router, a network switch, a network bridge, Personal Digital Assistant (PDA), a cellular telephone, a web appliance or any machine capable of executing a sequence of instructions that specify actions to be taken by

that machine.

[0069] The computer system 800 includes a processor 802, a main memory 804 and a static memory 806, which communicate with each other via a bus 808. The computer system 800 may further include a video display unit 810 (e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT)). The computer system 800 also includes an alpha-numeric input device 812 (e.g., a keyboard), a cursor control device 814 (e.g., a mouse), a disk drive unit 816, a signal generation device 820 (e.g., a speaker) and a network interface device 822.

[0070] The disk drive unit 816 includes a computer-readable medium 824 on which is stored a set of instructions (i.e., software) 826 embodying any one, or all, of the methodologies described above. The software 826 is also shown to reside, completely or at least partially, within the main memory 804 and/or within the processor 802. The software 826 may further be transmitted or received via the network interface device 822. For the purposes of this specification, the term " computer-readable medium" shall be taken to include any medium that is capable of storing or encoding a sequence of instructions for execution by the computer and that cause the computer to perform any one of the methodologies of the present invention. The term "computer-readable medium" shall accordingly be taken to included, but not be limited to, solid-state memories, optical and magnetic disks, and carrier wave signals.

[0071] Whereas many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after having read the foregoing description, it is to be understood that any particular embodiment shown and described by way of illustration is in no way intended to be considered limiting. Therefore,

references to details of various embodiments are not intended to limit the scope of the claims which in themselves recite only those features regarded as essential to the invention.

CLAIMS

We claim:

1. A method comprising:

receiving a plurality of probe email messages indicative of spam;

modifying the plurality of probe email messages to reduce noise;

comparing the plurality of probe email messages using fuzzy logic to identify similar email messages; and

creating groups of similar email messages, each of the groups pertaining to a distinct spam attack.

2. The method of claim 1 further comprising:

creating a filter for each of the groups; and

distributing resulting filters to a plurality of clients for detection of spam email messages at the plurality of clients.

3. The method of claim 1 wherein the plurality of probe email messages are collected using a plurality of fictitious probe email addresses selected to appear on spam email mailing lists.

4. The method of claim 1 wherein modifying the plurality of probe email messages to reduce noise comprises removing data indicative of noise from the plurality of probe email messages, the data indicative of noise being selected from the group

consisting of formatting data, numeric character references, character entity references, and uniform resource locator (URL) data of predefined categories.

5. The method of claim 1 wherein comparing the plurality of probe email messages comprises:

generating a mathematical signature of each of the plurality of probe email messages; and

finding similar mathematical signatures.

6. The method of claim 5 further comprising:

determining that a size of one of the plurality of probe email messages is below a threshold;

finding a predefined feature in the one of the plurality of probe email messages;

appending the predefined feature to the one of the plurality of probe email messages; and

if the size of the one of the plurality of probe email messages remains below the threshold, continuing to append the predefined feature to the one of the plurality of probe email messages until the size of one of the plurality of probe email messages reaches the threshold.

7. The method of claim 6 wherein the predefined feature is any one of a uniform resource locator (URL), a phone number, a keyword, and a name.

8. A method comprising:

determining that a size of a new message indicative of spam is below a threshold;

finding a predefined feature in the new message;

appending the predefined feature to the new message until the size of the new message reaches the threshold;

determining whether the new message is similar to a message from any of existing message groups; and

if the new message is different from messages in the existing message groups, creating a new message group denoting a new spam attack.

9. The method of claim 8 wherein the new message is a probe email message collected using one of a plurality of fictitious probe email addresses selected to appear on spam email mailing lists.

10. The method of claim 8 further comprising:

removing data indicative of noise from the new message prior to determining that the size of the new message is below the threshold, the data indicative of noise being selected from the group consisting of formatting data, numeric character references, character entity references, and uniform resource locator (URL) data of predefined categories.

11. The method of claim 8 wherein determining whether the new message is similar to messages from any of the existing message groups comprises:

generating a mathematical signature of the new message; and

determining whether the mathematical signature of the new message is similar to mathematical signatures of messages from any of the existing groups.

12. The method of claim 8 wherein the predefined feature is any one of a uniform resource locator (URL), a phone number, a keyword, and a name.

13. A system comprising:

probe email parser to receive a plurality of probe email messages indicative of spam;

noise reduction algorithm to modify the plurality of probe email messages to reduce noise; and

a grouping module to compare the plurality of probe email messages using fuzzy logic to identify similar email messages, and to create groups of similar email messages, each of the groups pertaining to a distinct spam attack.

14. The system of claim 13 wherein the grouping module comprises a signature generator to generate a mathematical signature of each of the plurality of probe email messages, and a signature matching algorithm to find similar mathematical signatures.

15. The system of claim 14 wherein the grouping module further comprises a feature extractor to determine that a size of one of the plurality of probe email messages is below a threshold, to find a predefined feature in the one of the plurality of probe email messages, to append the predefined feature to the one of the plurality of probe email

messages, and to continue appending the predefined feature to the one of the plurality of probe email messages until the size of one of the plurality of probe email messages reaches the threshold.

16. The system of claim 15 wherein the predefined feature is any one of a uniform resource locator (URL), a phone number, a keyword, and a name.

17. The system of claim 13 further comprising:

a filter generator to create a filter for each of the groups; and

a filter transmitter to distribute resulting filters to a plurality of clients for detection of spam email messages at the plurality of clients.

18. The system of claim 13 wherein the plurality of probe email messages are collected using a plurality of fictitious probe email addresses selected to appear on spam email mailing lists.

19. An apparatus comprising:

means for receiving a plurality of probe email messages indicative of spam;

means for modifying the plurality of probe email messages to reduce noise;

means for comparing the plurality of probe email messages using fuzzy logic to identify similar email messages; and

means for creating groups of similar email messages, each of the groups pertaining to a distinct spam attack.

20. A computer readable medium comprising executable instructions which when executed on a processing system cause said processing system to perform a method comprising:

receiving a plurality of probe email messages indicative of spam;

modifying the plurality of probe email messages to reduce noise;

comparing the plurality of probe email messages using fuzzy logic to identify similar email messages; and

creating groups of similar email messages, each of the groups pertaining to a distinct spam attack.

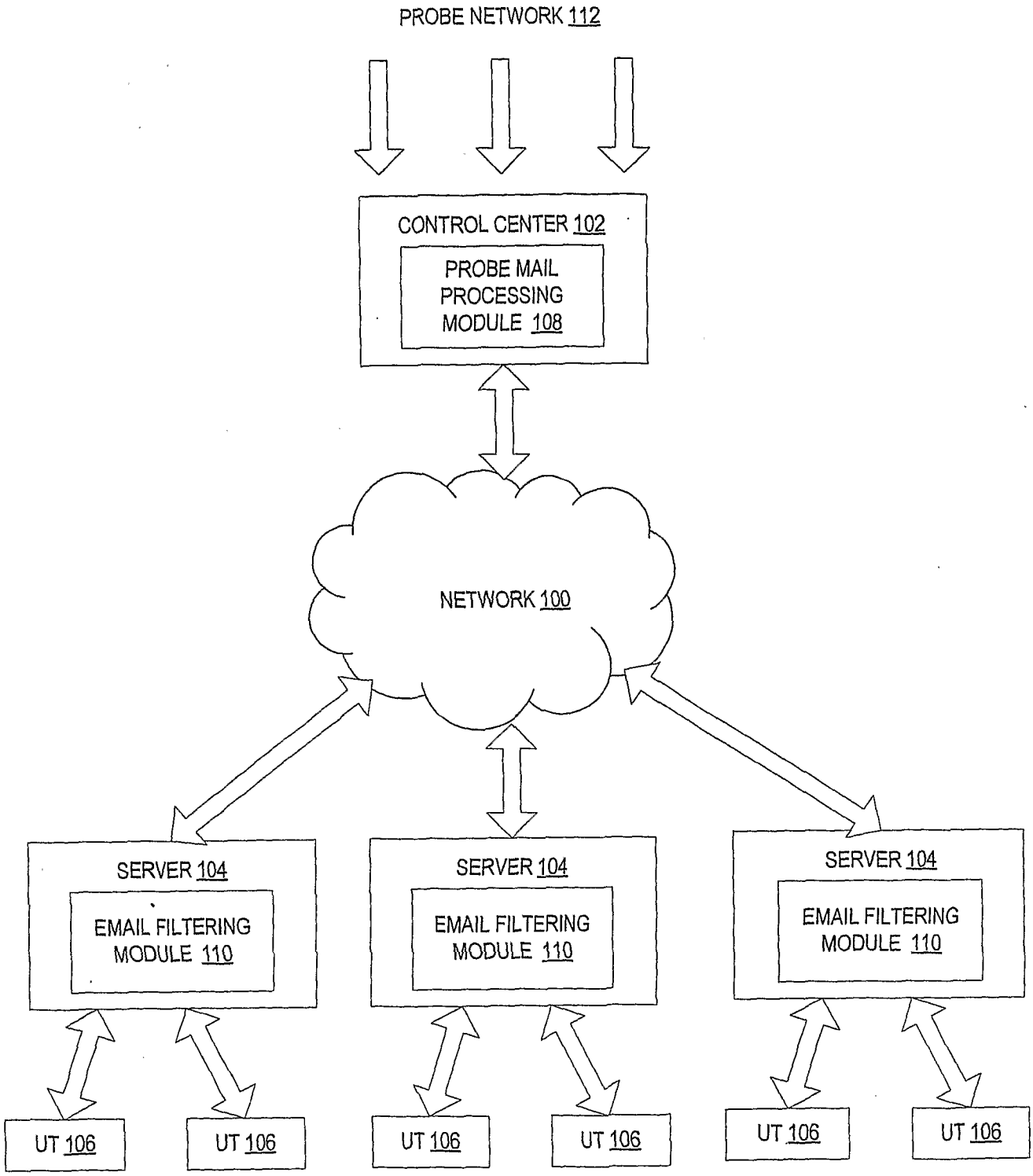

PROBE NETWORK 112
CONTROL CENTER 102
PROBE MAIL PROCESSING MODULE 108
NETWORK 100
SERVER 104
EMAIL FILTERING MODULE 110
SERVER 104
EMAIL FILTERING MODULE 110
SERVER 104
EMAIL FILTERING MODULE 110
UT 106
UT 106
UT 106
UT 106
UT 106
UT 106

FIG. 1

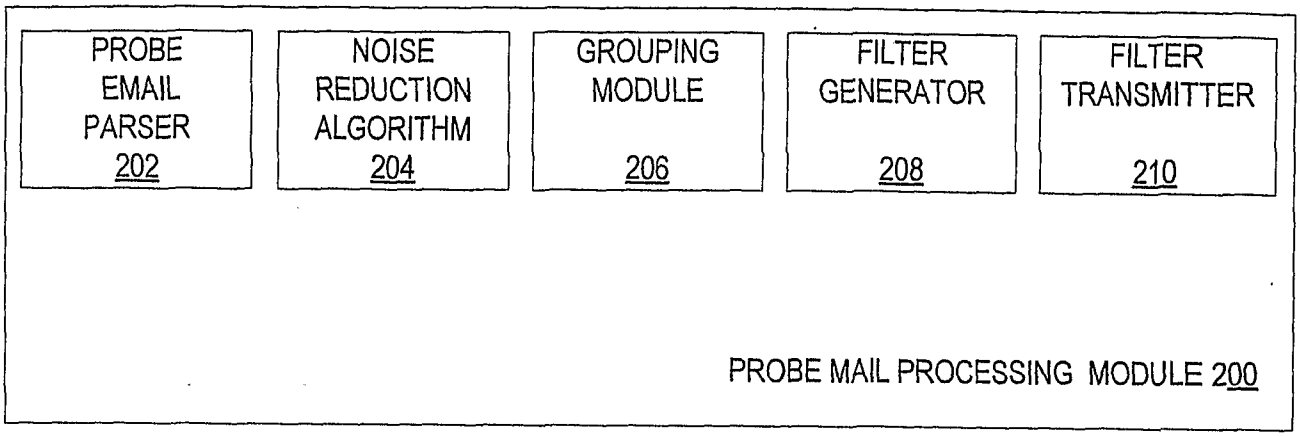

PROBE EMAIL PARSER 202
NOISE REDUCTION ALGORITHM 204
GROUPING MODULE 206
FILTER GENERATOR 208
FILTER TRANSMITTER 210
PROBE MAIL PROCESSING MODULE 200

FIG. 2

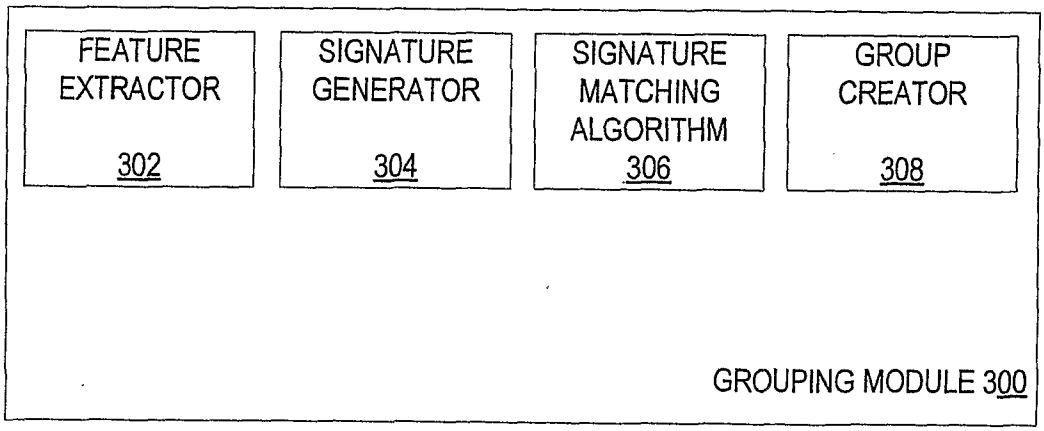

FEATURE EXTRACTOR 302
SIGNATURE GENERATOR 304
SIGNATURE MATCHING ALGORITHM 306
GROUP CREATOR 308
GROUPING MODULE 300

FIG. 3

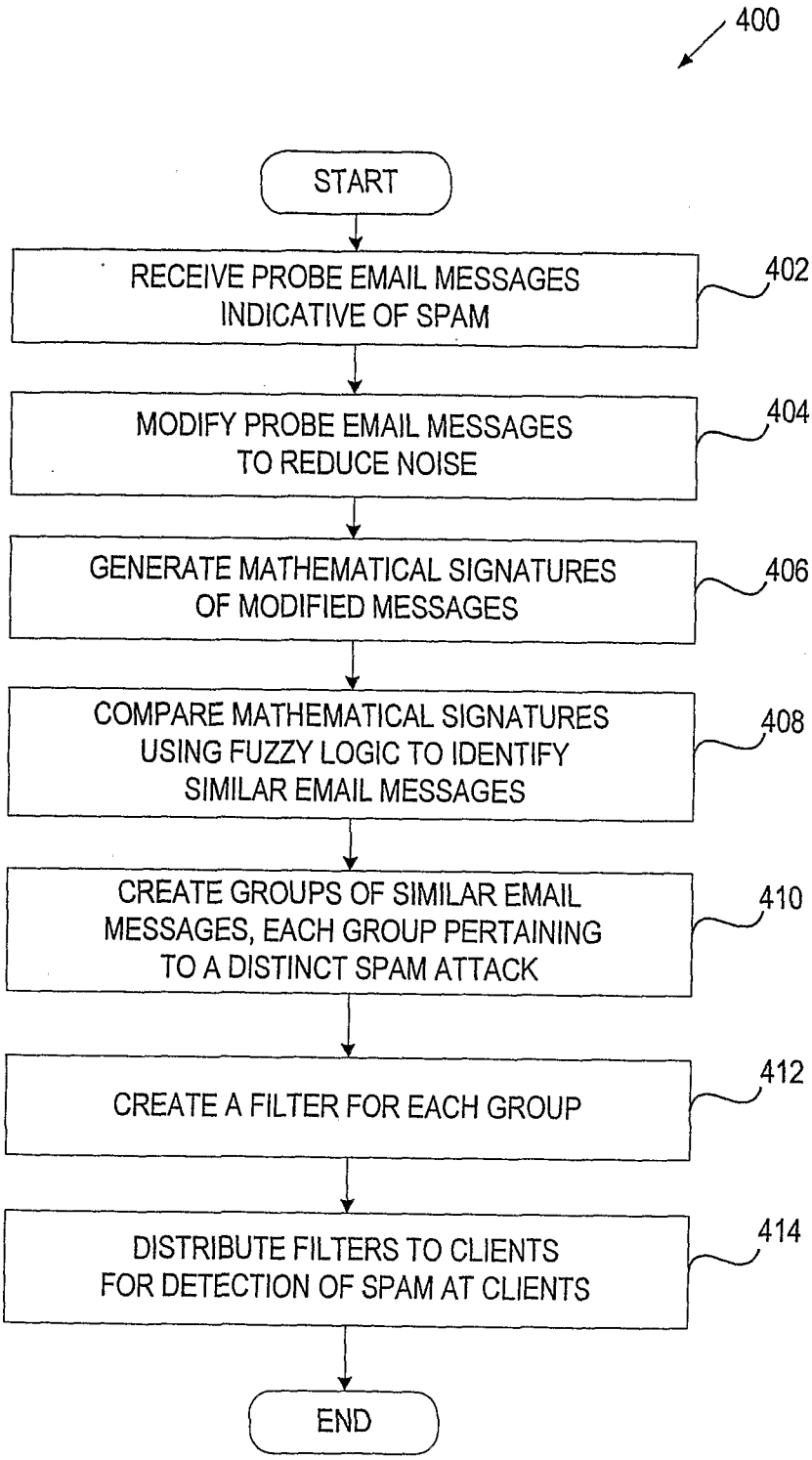

400
START
RECEIVE PROBE EMAIL MESSAGES INDICATIVE OF SPAM
402
MODIFY PROBE EMAIL MESSAGES TO REDUCE NOISE
404
GENERATE MATHEMATICAL SIGNATURES OF MODIFIED MESSAGES
406
COMPARE MATHEMATICAL SIGNATURES USING FUZZY LOGIC TO IDENTIFY SIMILAR EMAIL MESSAGES
408
CREATE GROUPS OF SIMILAR EMAIL MESSAGES, EACH GROUP PERTAINING TO A DISTINCT SPAM ATTACK
410
CREATE A FILTER FOR EACH GROUP
412
DISTRIBUTE FILTERS TO CLIENTS FOR DETECTION OF SPAM AT CLIENTS
414
END

FIG. 4

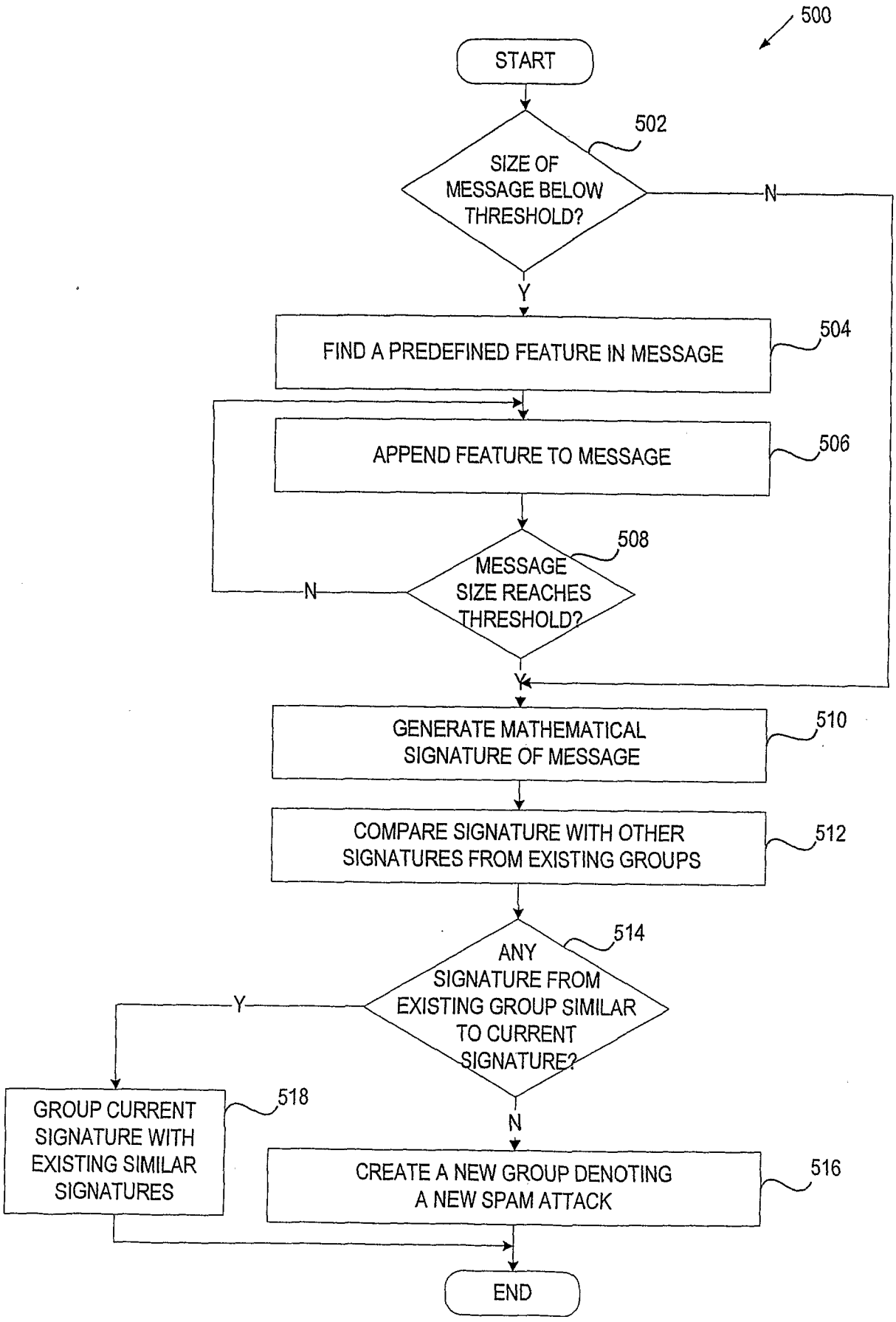
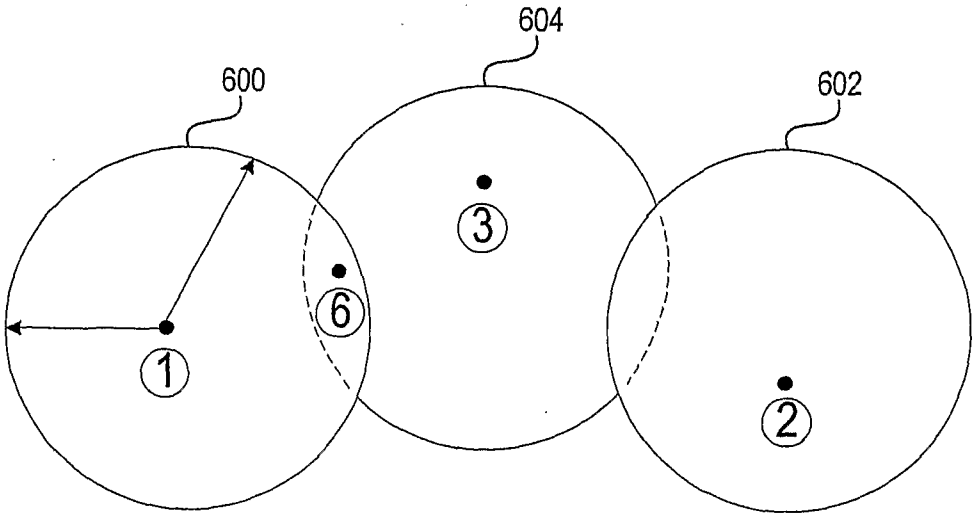

500
START
502
SIZE OF MESSAGE BELOW THRESHOLD?
N
Y
FIND A PREDEFINED FEATURE IN MESSAGE
504
APPEND FEATURE TO MESSAGE
506
508
MESSAGE SIZE REACHES THRESHOLD?
N
Y
GENERATE MATHEMATICAL SIGNATURE OF MESSAGE
510
COMPARE SIGNATURE WITH OTHER SIGNATURES FROM EXISTING GROUPS
512
514
ANY SIGNATURE FROM EXISTING GROUP SIMILAR TO CURRENT SIGNATURE?
Y
N
GROUP CURRENT SIGNATURE WITH EXISTING SIMILAR SIGNATURES
518
CREATE A NEW GROUP DENOTING A NEW SPAM ATTACK
516
END

FIG. 5



1 = > FIRST ATTACK
2 = > SECOND ATTACK
3 = > NOT IN 1 OR 2 BUT OVERLAPS
6 = > IN 1 AND 3

FIG. 6

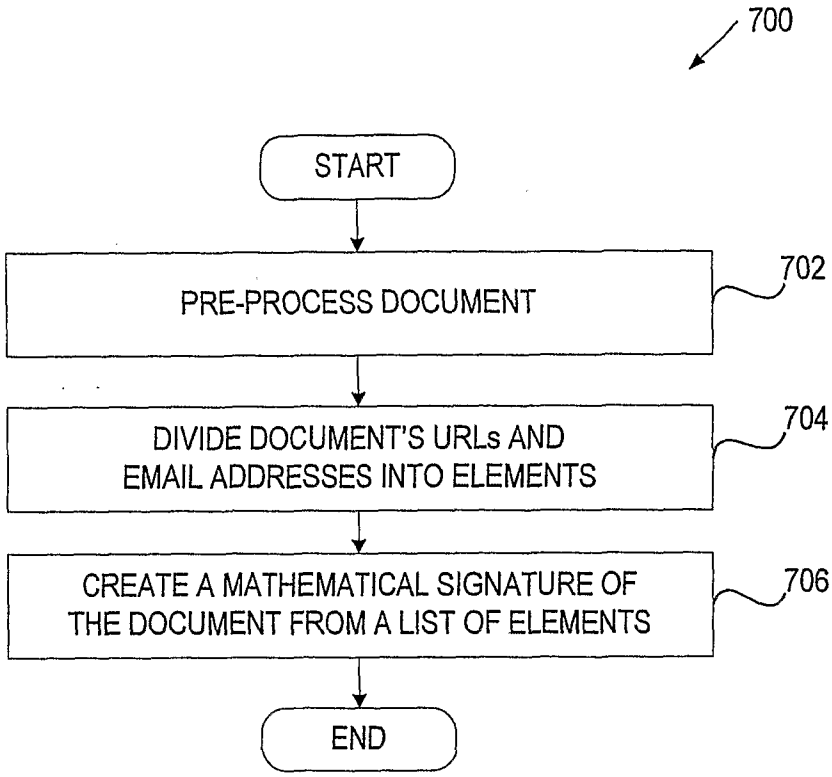

700
START
PRE-PROCESS DOCUMENT
702
DIVIDE DOCUMENT'S URLs AND EMAIL ADDRESSES INTO ELEMENTS
704
CREATE A MATHEMATICAL SIGNATURE OF THE DOCUMENT FROM A LIST OF ELEMENTS
706
END

FIG. 7

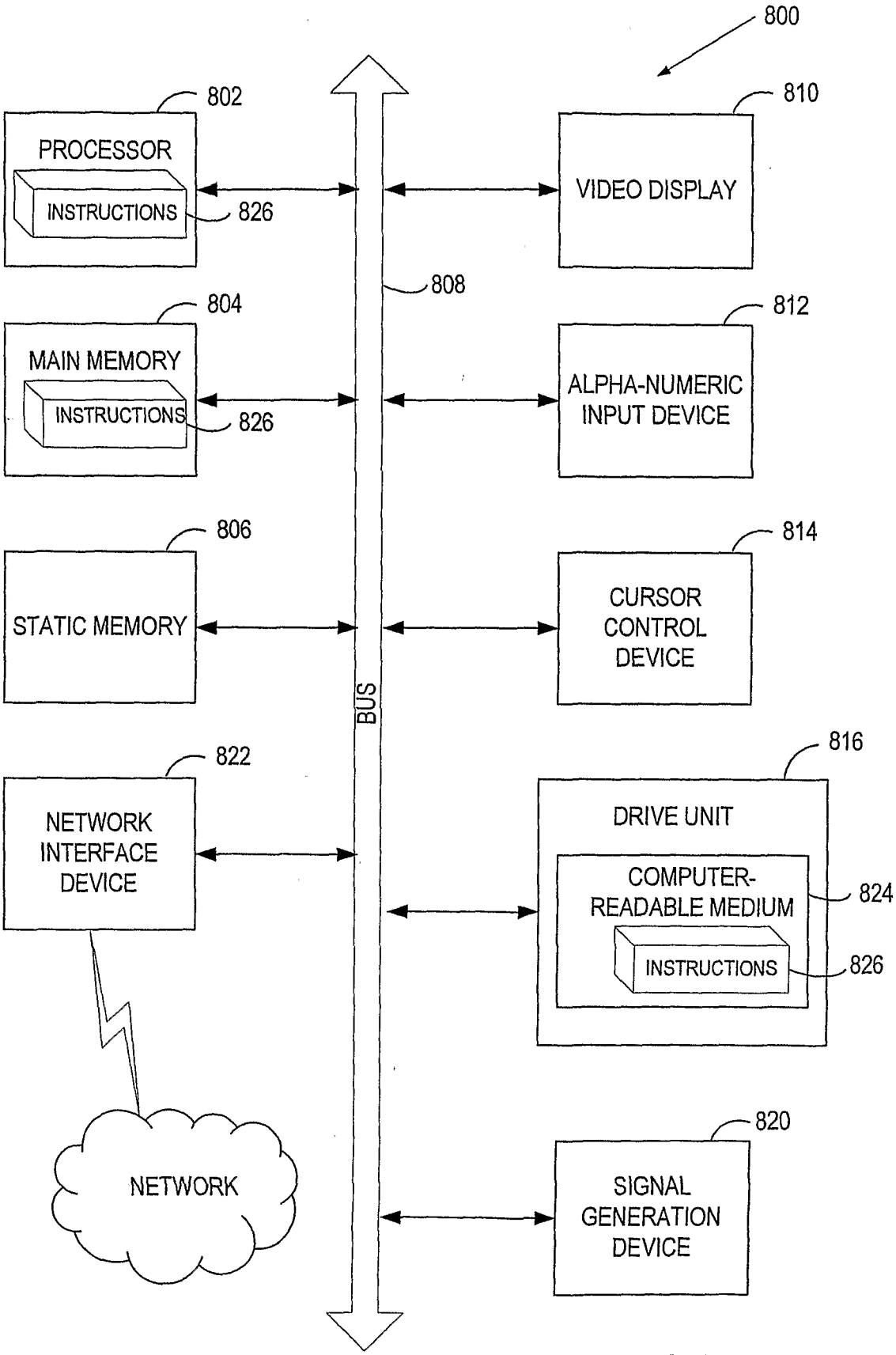

800
802
PROCESSOR
INSTRUCTIONS
826
804
MAIN MEMORY
INSTRUCTIONS
826
806
STATIC MEMORY
822
NETWORK INTERFACE DEVICE
NETWORK
808
BUS
810
VIDEO DISPLAY
812
ALPHA-NUMERIC INPUT DEVICE
814
CURSOR CONTROL DEVICE
816
DRIVE UNIT
COMPUTER-READABLE MEDIUM
824
INSTRUCTIONS
826
820
SIGNAL GENERATION DEVICE

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2006/023847

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04L12/58

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

B. FIELDS SEARCHED

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

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 01/53965 A (ODYSSEY DEVELOPMENT PTY LTD; DAVIES, IAN, ALEXANDER) 26 July 2001 (2001-07-26) page 5, line 17 - page 6, line 9 page 8, line 24 - page 9, line 33 page 11, line 1 - line 19 claim 1	1-5,13, 14,17-20 6-12,15, 16
Y A	US 2005/060643 A1 (GLASS JEFFREY BRIAN ET AL) 17 March 2005 (2005-03-17) paragraphs [0061], [0065], [0066], [0104] - [0125]	1-5,13, 14,17-20 6-12,15, 16
A	RUBIN P: "Re: spam (hash codes)" RE: SPAM (HASH CODES), 20 February 1995 (1995-02-20), XP002229452 the whole document	1-20

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 October 2006	10/10/2006

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

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2006/023847

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 052 709 A (PAUL ET AL) 18 April 2000 (2000-04-18) column 3, line 46 - column 5, line 32 column 7, line 37 - line 54 figures 1,8	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2006/023847

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 0153965	A	26-07-2001	NONE		
US 2005060643	A1	17-03-2005	NONE		
US 6052709	A	18-04-2000	AT	310989 T	15-12-2005
			AU	1631199 A	12-07-1999
			DE	69832524 D1	29-12-2005
			DE	69832524 T2	10-08-2006
			EP	1040584 A2	04-10-2000
			ES	2252873 T3	16-05-2006
			HK	1031956 A1	28-07-2006
			JP	2001527257 T	25-12-2001
			WO	9933188 A2	01-07-1999