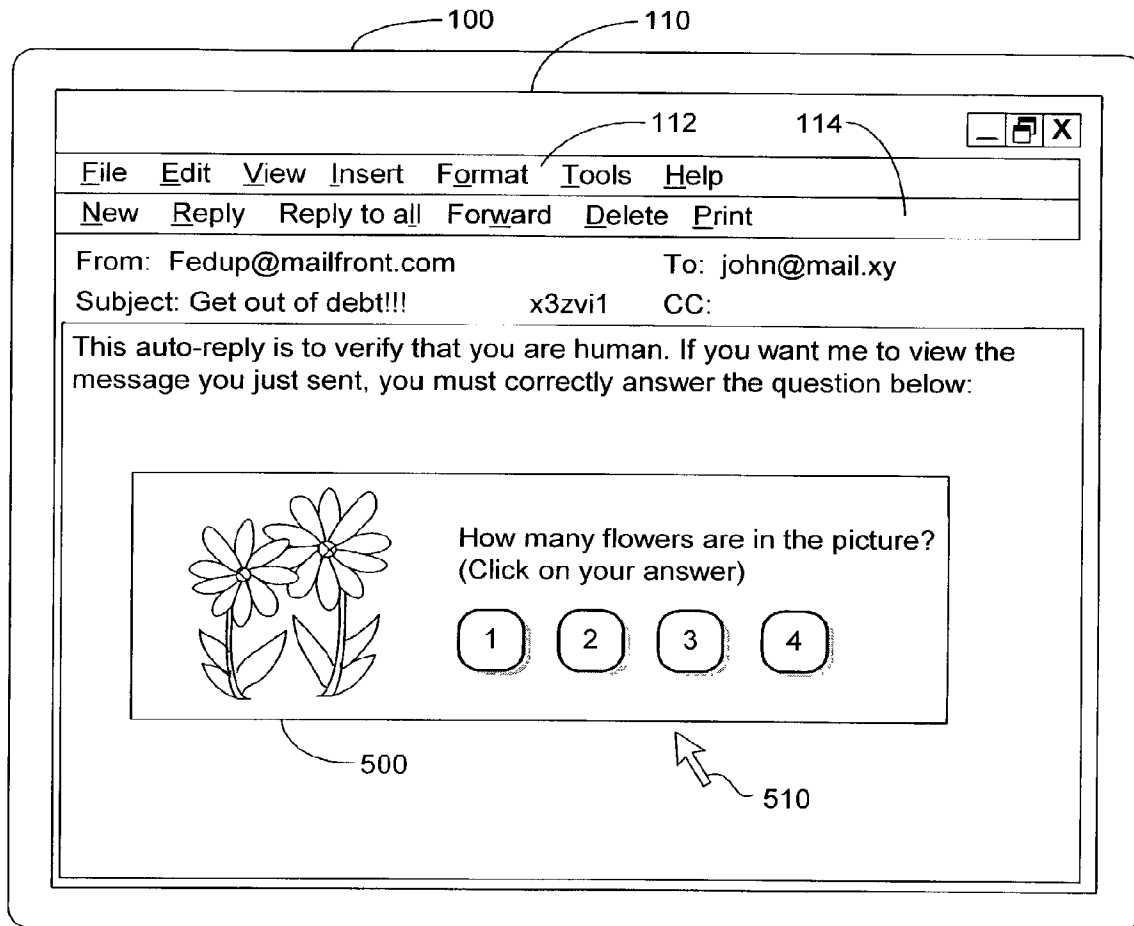




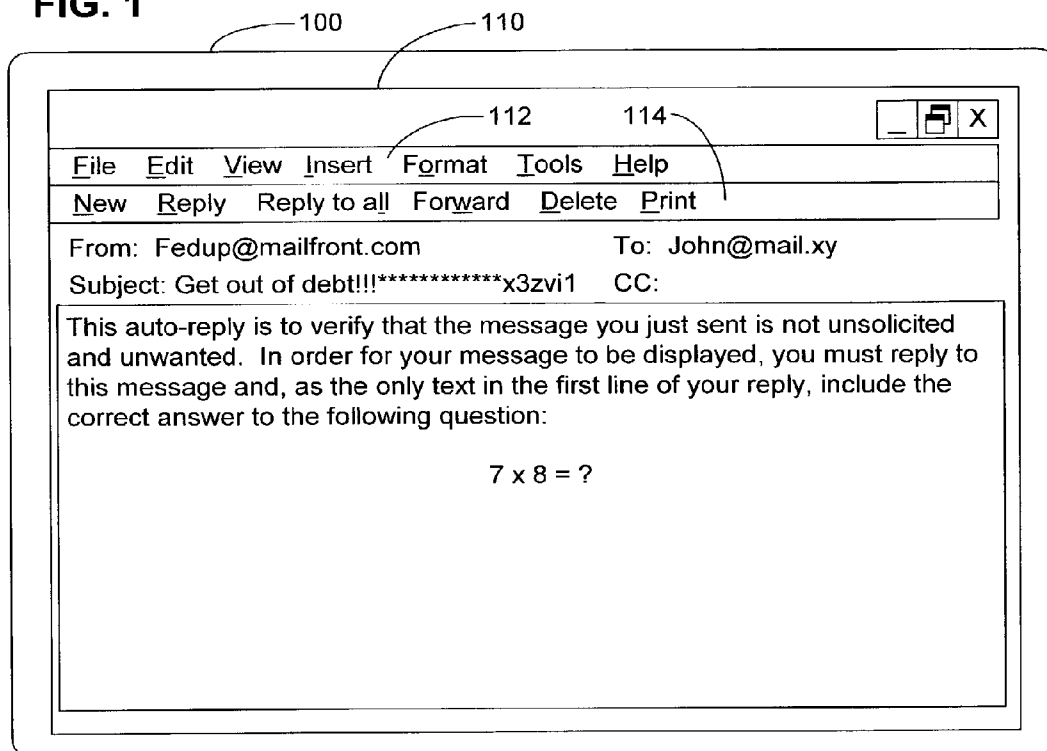
US 20080168145A1

(19) **United States**(12) **Patent Application Publication**  
**Wilson**(10) **Pub. No.: US 2008/0168145 A1**(43) **Pub. Date: Jul. 10, 2008**(54) **ACTIVE E-MAIL FILTER WITH  
CHALLENGE-RESPONSE**(52) **U.S. Cl. .... 709/206**(76) **Inventor: Brian Wilson, Palo Alto, CA (US)****Correspondence Address:**  
**CARR & FERRELL LLP**  
**2200 GENG ROAD**  
**PALO ALTO, CA 94303**(57) **ABSTRACT**

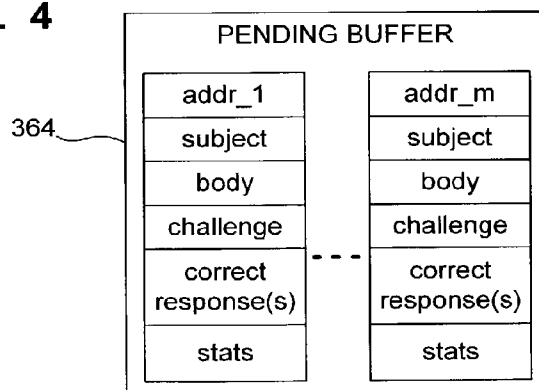
When a user receives potential junk e-mail from an unknown sender address, an active filter installed in or cooperating with the user's e-mail system sends a challenge e-mail back to the sender address. Unless the user receives a correct response e-mail from the sender address and/or a response that meets formal criteria, the active filter rejects the original e-mail and blocks future e-mail from that sender address. The challenge preferably includes a question that the sender must correctly answer in his response. The question is such that only humans can interpret and respond to it correctly. For example, the question may relate to and/or be incorporated in an image in the challenge. Lists of approved and blocked sender addresses are compiled. The user may directly enter addresses in the address lists. Addresses to which the user has sent e-mail may also be included directly in the approved list.

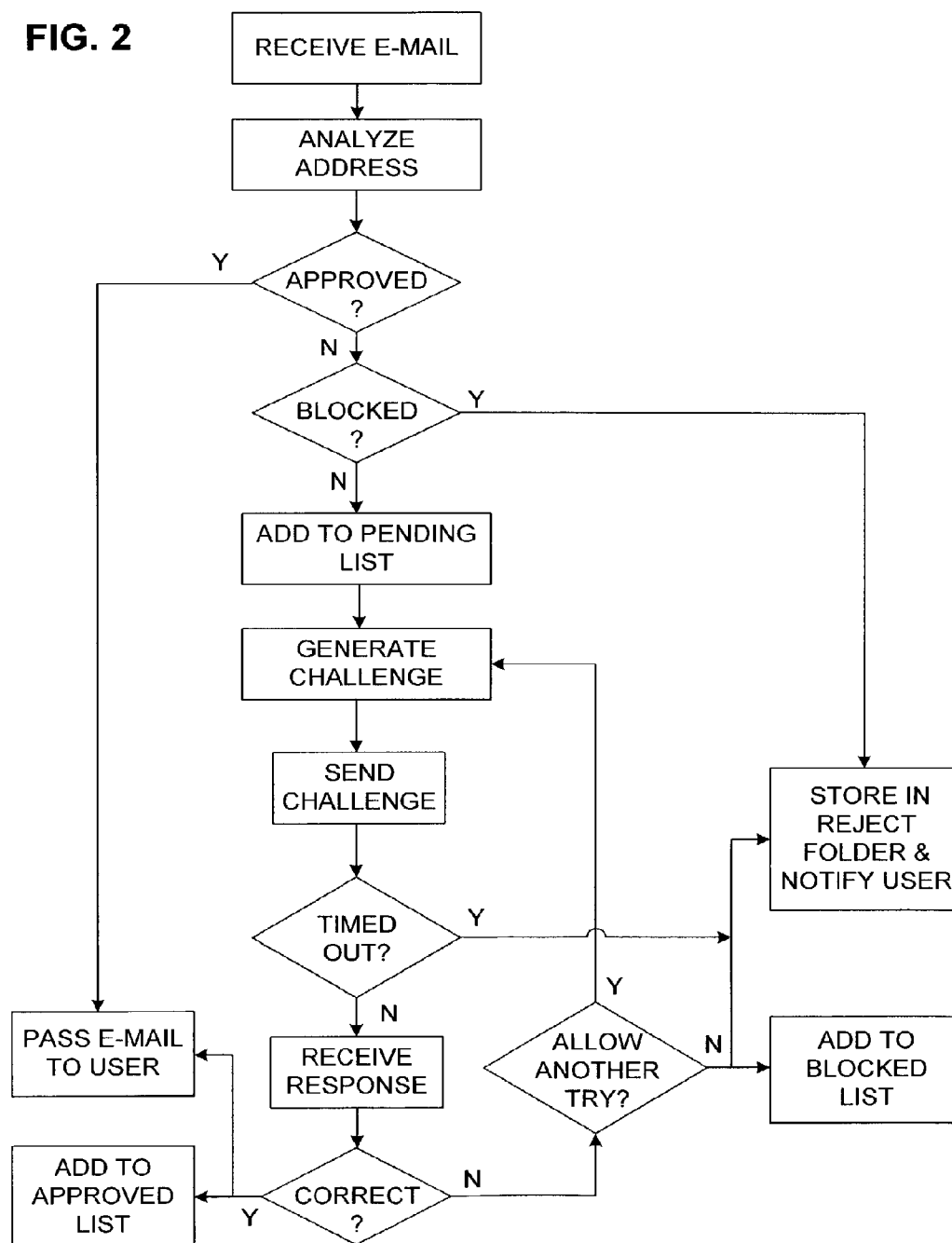
(21) **Appl. No.: 11/927,414**(22) **Filed: Oct. 29, 2007****Related U.S. Application Data**(63) **Continuation of application No. 10/197,393, filed on  
Jul. 16, 2002.****Publication Classification**(51) **Int. Cl.**  
**G06F 15/16 (2006.01)**

**FIG. 1**

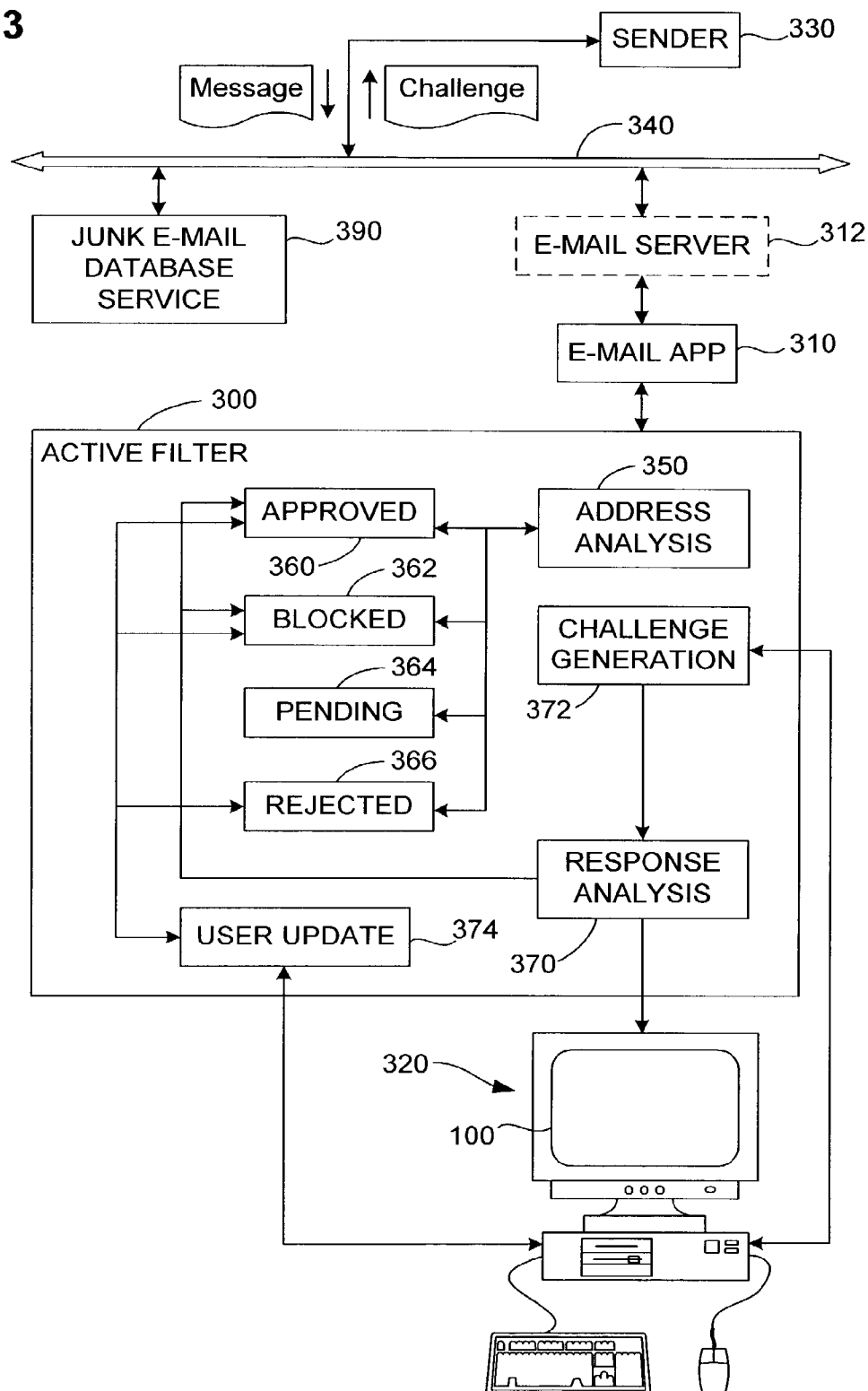


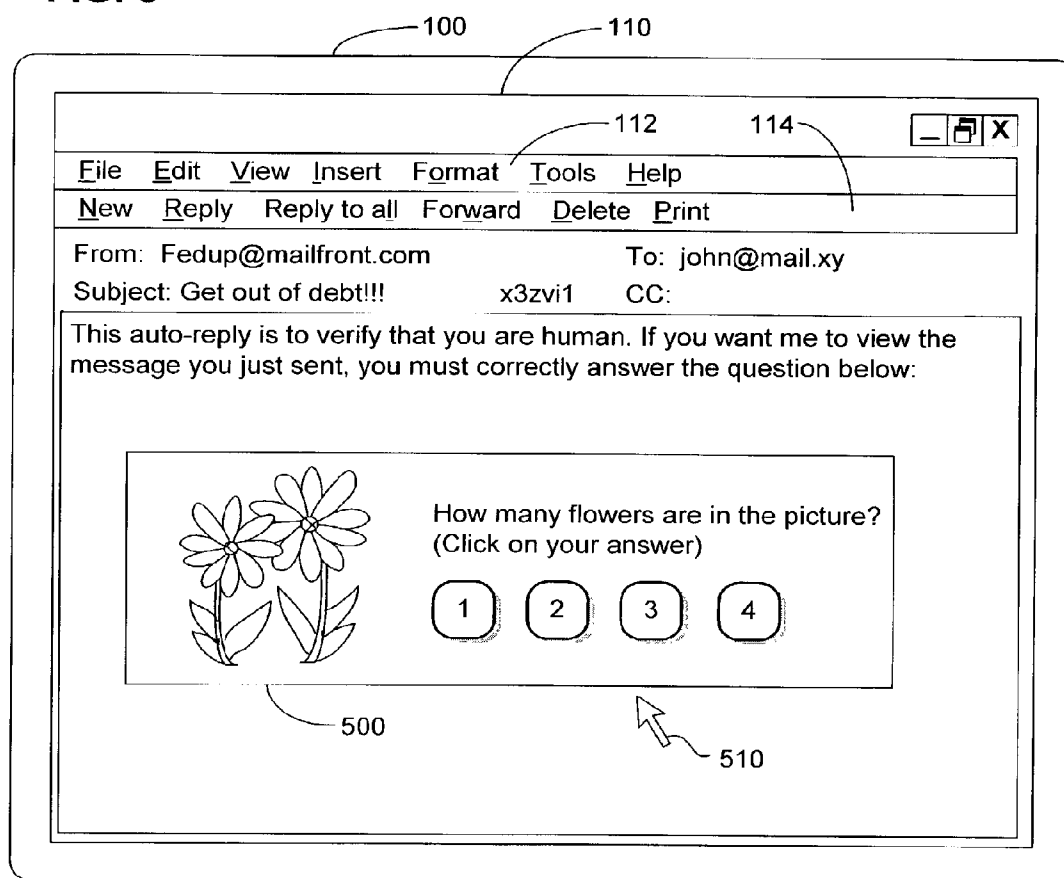
**FIG. 4**



**FIG. 2**

**FIG. 3**



**FIG. 5**

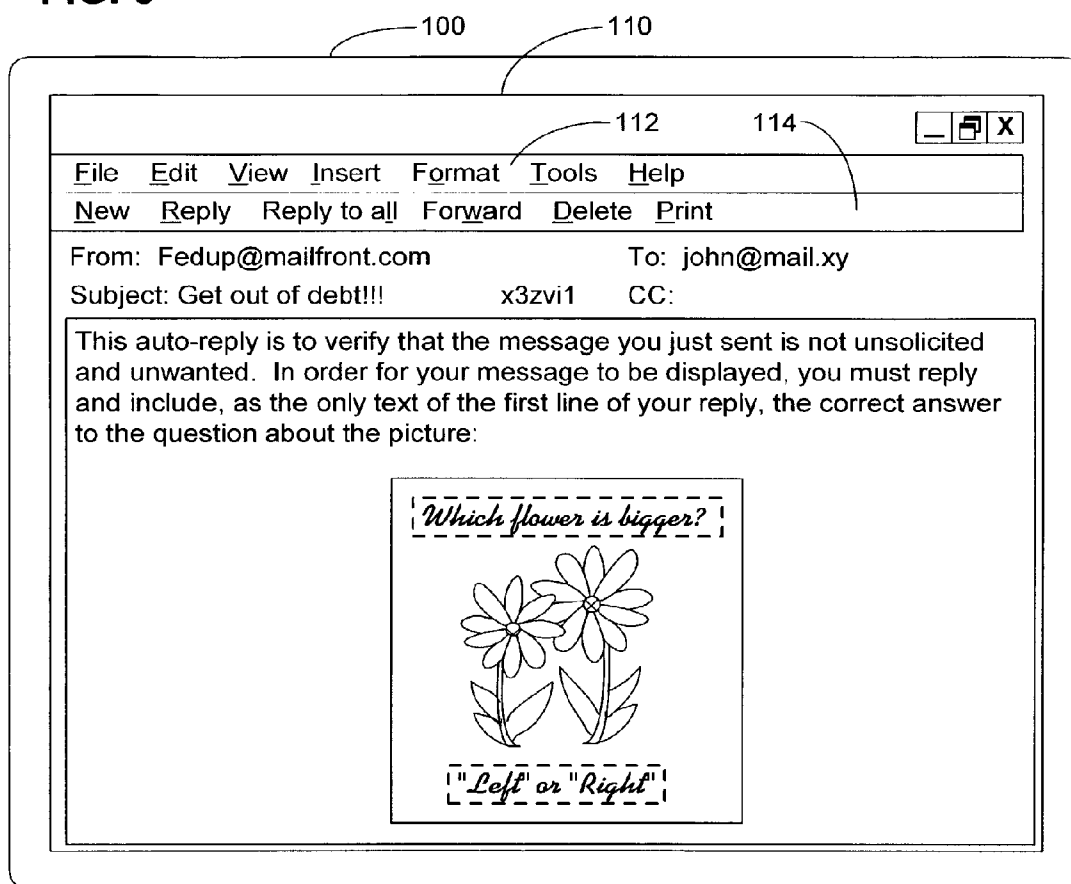
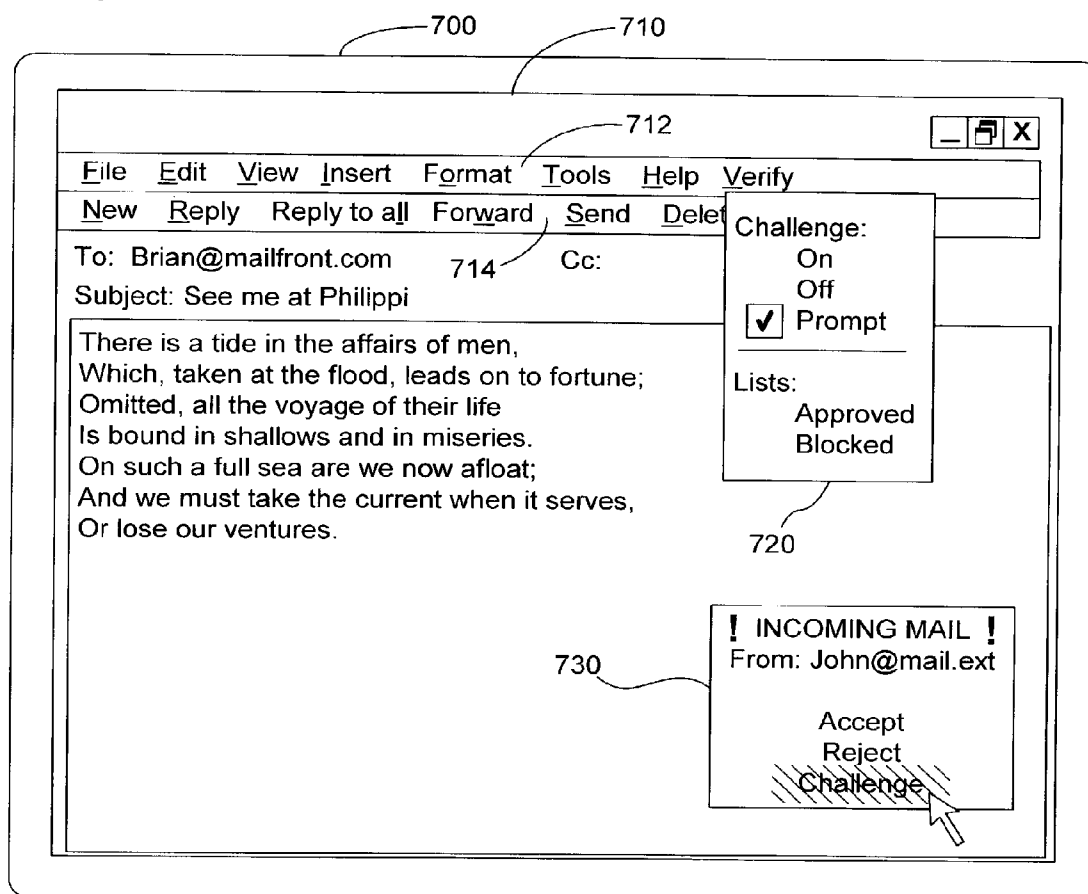
**FIG. 6**

FIG. 7



## ACTIVE E-MAIL FILTER WITH CHALLENGE-RESPONSE

### BACKGROUND OF THE INVENTION

**[0001]** 1. Field of the Invention

**[0002]** This invention relates to a system and method for eliminating or at least reducing the amount of unwanted electronic mail messages sent to users of e-mail.

**[0003]** 2. Description of Related Art

**[0004]** As the use of electronic mail (e-mail) has grown, so too has the amount of unwanted e-mail. This unwanted e-mail, commonly referred to as "junk e-mail" or "spam," is usually, but not always, some form of solicitation to buy some product or service.

**[0005]** Junk e-mail in the "Inbox" of e-mail programs is analogous to the "junk mail" that fills many mail boxes. The analogy is not perfect, however. Perhaps the greatest difference, besides the obvious difference in media, is that the marginal cost of sending junk e-mail to one more recipient is essentially nil, whereas an address label must be printed and postage must be paid for each additional piece of junk mail sent through normal postal channels. Indeed, the alarming growth of the junk e-mail industry depends on this cost efficiency. Even millions of e-mail addresses can be obtained relatively cheaply, especially since many organizations that have compiled such lists for other reasons are often happy to get the "free" money that is to be had simply by selling them. Even absent such straightforward purchase, less scrupulous "junk e-mailers" can still get at addresses by using known computer programs that gain access to and download address lists from servers. Even more disturbingly, junk e-mailers can also launch programs that cause central e-mail servers themselves to propagate the junk e-mail to their own customers.

**[0006]** That junk e-mail is almost universally hated is of little concern to its creators, once again because of the negligible cost: Assume that 9,999 out of 10,000 who receive an e-mail solicitation for a \$100 service are irate at having received it at all, but that one actually buys the service. If the junk e-mailer manages to transmit just a million junk e-mails with the solicitation, then he will still have made \$10,000, which is typically far more than it would cost to generate the solicitation. Moreover, note that there are even on-line companies that specialize in sending junk e-mail on behalf of others, usually for only a few hundred dollars, thereby reducing the cost of sending mass junk e-mail even further. Even these companies send junk e-mail advertising their service, which constitutes junk e-mail that advertises a service to enable others to send junk e-mail!

**[0007]** It is not just countless recipients of junk e-mail who would like to rid the e-mail system of it: The hundreds of millions of junk e-mail messages sent every day (and perhaps every hour or minute) represent a waste of the bandwidth of the network (usually Internet) and thus a loss to e-mail service providers as well.

**[0008]** There are at present many different systems that attempt to filter out junk e-mail. Some of these systems are incorporated in the user's e-mail program itself, others are installed in the central e-mail server, and still others are separate applications that the user must load into his computer.

**[0009]** Using perhaps the most common system, when a user receives unwanted e-mail, he may direct his e-mail application to put the sender's address in a "blocked" list. Future messages from the same address are then prevented from reaching the user's Inbox, and are instead sometimes routed

to a special junk e-mail folder for optional review by the user. The weakness of this filtering method is that junk e-mailers can easily change the address but keep the same domain name. The junk e-mail will then pass this simple filter. Thus, the sender address "aaa@junk.ext" might be blocked, but "aab@junk.ext" will not be.

**[0010]** One improvement over simple address filtering is therefore to filter based on domain names. There are two weaknesses to this approach. First, it is possible to change the "Reply to" or "From" address of a sent e-mail message. Automatically generated, fake domain names can therefore be sent along with junk e-mail to bypass domain name-based filters. Second, junk e-mailers often misappropriate common domain names, especially if the junk e-mail is being sent by a "Trojan Horse" program from within a widely used e-mail server; users may not want to block all e-mail from that domain.

**[0011]** Other known junk e-mail filters examine the routing information that accompanies each e-mail in an attempt to identify the original address from which the e-mail was sent. This will still not work when the e-mail is sent from within a central "trusted" server. Even so, junk e-mailers may also use known techniques to cause their messages to be sent through different proxy servers and thereby to disguise their origin.

**[0012]** The "Active SMTP" technology of Escom Corp. represents another approach, according to which a sending e-mail host is probed while it is connected to the recipient's server. If the sending host is a dial-up risk, or an open relay, or is configured not to accept incoming e-mail (typically indicative of a forged address), then the Active SMTP system distrusts the sending host and rejects its e-mail. This system has at least two specific weaknesses: First, it will not block unwanted e-mail such as blatant commercial advertisements if the sender is using regular e-mail sent from a properly configured mail server. Second, this system works only when the sending server is connected to the recipient server; consequently, such a system is not suitable for installation on users' local computer systems that download e-mail from a remote server.

**[0013]** Still other known junk e-mail filters examine not just addresses and routings, but rather (or in addition) some aspect of the content of the messages. Content is usually found either in the "subject line" of the message, or the main text of the message, or both. It may also be found in the message header. Assume such a filter is installed in a central e-mail server. If a filter program notices that a million pending e-mail messages all have the same subject line, or have message texts with identical wording, then it is almost certain that these messages are junk e-mail and can be blocked and deleted. To defeat such filters, junk e-mail senders typically add random strings of characters in the subject line (usually after many blank spaces or asterisks to make them less obvious to recipients) or at the bottom of the texts. Character-by-character or word-by-word comparison of any two of the junk messages will therefore show them to be different, although the only difference is the meaningless, random character strings included specifically to defeat the filters.

**[0014]** Other "tricks" used by senders of junk e-mail include fake personalization of the messages. Thus, by including all or some part of the recipient's name (extracted from his e-mail address or from a bulk address list) in the subject line, the sender attempts to fool the recipient into



believing that the message is legitimate. For example, the text “Bob, here is the deal I told you about!” might be in the subject line of a message sent to Bob@mailfront.com. This ruse alone may defeat some known filters and may induce the unwitting recipient to open the message. This trick of false familiarity is of course also well known from junk mail sent by post. Note that many of the methods used by junk e-mailers resemble those used to transmit computer viruses, which often require that recipients be tricked into opening some executable attachment.

**[0015]** Legislation has done little to combat the plague of junk e-mail. Indeed, there is proposed federal legislation against “unsolicited commercial e-mail,” but it lacks provision for effective enforcement and has loopholes through which junk e-mailers easily slip. Like laws, evasion is often a simple exercise in semantics. For example, many junk e-mail messages include false or at best misleading statements indicating that they are not unsolicited, or not commercial. Some particularly brazen junk e-mail even states explicitly that it complies with the very proposed legislation it violates. All such statements are of course intended to confuse recipients or at least make them believe they have no recourse. Examples of actual statements received by the applicant’s attorney include:

**[0016]** “Below is the result of your feedback form.” (No feedback form was ever submitted.);

**[0017]** “Note: this is not a spam email. This email was sent to you because your email was entered in on (sic) one of our websites requesting to be a registered subscriber.” (No such request had ever been submitted.)

**[0018]** “THIS IS NOT AN OFFER OR CONTRACT TO BUY INSURANCE PRODUCTS, but rather a confidential informational inquiry” (attempting to circumvent the prohibition against unsolicited commercial e-mail); and

**[0019]** “This message is being sent in compliance of (sic) the email bill: Section 301.per section, paragraph (a) (2) (c) of S. 1618.”

**[0020]** Still other junk e-mail filters work by searching the message for certain predetermined, prohibited keywords. Any messages that contain any of the keywords are blocked and are not passed through to the user’s e-mail application. These filters thus operate in the same manner as (and often are) known filters designed to eliminate pornography or other objectionable subject matter. The problems of over- and under-inclusiveness of these systems are well documented.

**[0021]** The general weakness of all of the known junk e-mail filters mentioned above stems from the fact that they are passive: They attempt to determine whether a message is junk e-mail based solely on an analysis of information extracted from and relating to the message and its sending server. As such, the senders of junk e-mail, like car thieves working to defeat alarms, can evolve their techniques so as to defeat the latest filters.

**[0022]** What is needed is a mechanism for eliminating or at least greatly reducing the successful transmission of unwanted e-mail while still making it easy and convenient to receive wanted e-mail. In other words, e-mail from unwanted senders should be effectively blocked, with as little interference as possible with e-mail from approved senders. It should

also be impossible—or at least exceptionally difficult—for junk e-mailers to bypass the filter. This invention provides such a mechanism.

## SUMMARY OF THE INVENTION

**[0023]** The invention provides a method and system implementation for filtering electronic mail (e-mail) according to which, when a current e-mail message is received from a sender address, a challenge that requires a correct e-mail response is returned, also by e-mail, to the sender address. In the simplest embodiment of the invention, the current e-mail message from the sender address is then rejected unless the correct e-mail response is received by the user system from the sender address. All future e-mail messages from the sender address are preferably also blocked.

**[0024]** The challenge includes an instruction to perform an action; the correct e-mail response requires correct performance of the action. The challenge is preferably in the form of a question that must be correctly answered, and preferably includes an element that only a human can correctly interpret, such as an image, audio clip, text question whose answer requires human experience, etc. The challenge is preferably generated automatically by a corresponding software module, which may alternatively download challenges from an external system.

**[0025]** Challenge responses that do not meet formal criteria are preferably also rejected. Such formal criteria include response within a maximum allowed time and answering correctly within a maximum number of tries.

**[0026]** When a correct e-mail response is received from the sender address, that sender address is preferably included in a list of approved sender addresses and the current e-mail message may be accepted, meaning that it is passed for display to and review by the user, for example in an Inbox. The user may optionally directly specify e-mail addresses to be included in the list of approved sender addresses. Other e-mail addresses that may be assumed to be welcome and may be included directly in the approved list are those to which the user himself has initiated sending e-mail and addresses obtained from an externally compiled approved list.

**[0027]** When an incorrect e-mail response is received from the sender address, that sender address is preferably included in a list of blocked sender addresses. The current and future e-mail messages from the current sender address, or from any other address in the blocked list, are then rejected without further challenge. The user may also optionally directly specify e-mail addresses to be included in the blocked list; other blocked addresses from an externally compiled list may also be included. The list of blocked addresses may optionally be exported to other users.

**[0028]** Rather than automatically issuing challenges whenever e-mail is received from an unknown address, the system may instead notify the user of receipt of e-mail messages from such addresses and prompt him to indicate whether to send a challenge, to reject the message immediately, or to accept it without challenge.

**[0029]** For each challenge, a key is preferably calculated and included in the challenge. Received e-mail that includes a correctly calculated key may then be accepted without challenge (since it itself should be a challenge), or accepted

directly if the key is correct and the message is from a sender address to which the user previously initiated sending a message.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0030]** FIG. 1 illustrates an example of a text-based challenge to an e-mail sender.

**[0031]** FIG. 2 is a flowchart of the main steps performed by the active, challenge-response e-mail filter according to the invention.

**[0032]** FIG. 3 is a block diagram of the main hardware and software components used in the e-mail filter system according to the invention.

**[0033]** FIG. 4 illustrates an example of the structure of a buffer used to store information about e-mail messages whose senders are being challenged to prove they are human.

**[0034]** FIG. 5 illustrates an image-based challenge with an automatically launched, one-click response.

**[0035]** FIG. 6 illustrates an image-based challenge that requires a manually typed textual response.

**[0036]** FIG. 7 illustrates an example of a user interface that enables a user to turn the invention on and off and to activate prompted message challenging.

#### DETAILED DESCRIPTION

**[0037]** In broadest terms, the invention provides a mechanism for filtering e-mail messages that eliminates or at least greatly reduces the number of unwanted messages received. Unlike existing, passive filters, the invention achieves this by requiring the sender of questionable e-mail to prove that he is in fact human and not an automated mass-mailing system: Upon receipt of a message of unknown origin, an active filter in or associated with the recipient's e-mail system returns a challenge to the sender, which the sender must correctly respond to before the original message is actually passed on to the user. The challenge will typically be a question sent in a reply e-mail message, which the sender must correctly answer, also by e-mail. The filtering mechanism according to the invention is therefore active in that it puts the "burden of proof" on the sender rather than simply trying to passively identify mass or similar unsolicited e-mailings based solely on analysis of the information received.

**[0038]** By now, almost everyone in every industrialized country is very familiar with the structure and handling of an e-mail message. For the sake of completeness, however, this is summarized here: The "title" of an e-mail message is usually put in a "subject line," which the user's e-mail program displays along with the sender's address, the time and date the message was received, the size of the message, and possibly other identifying and general information, usually on a single line. As part of the standard protocol, an e-mail message also includes information that is not usually displayed, at least not by default. One example of such included but non-displayed information is a header.

**[0039]** If the user wants to view the main text of the message, he clicks on the subject line and the text is displayed on a separate part of the display screen. The message can then be deleted, forwarded, replied to, etc., by clicking on corresponding on-screen icons. Messages typically are first displayed in an Inbox folder, but they can be moved to other folders, usually by simple dragging and dropping using a mouse. Received messages can also be grouped automatically into different folders upon receipt; for example, differ-

ent users of the same e-mail program can have their own, separate default folders in which as yet unread e-mail is first placed.

**[0040]** Another common feature of modern e-mail programs is that they can automatically reply to received messages. This feature is often used to notify senders that the recipient is out of the office or otherwise unable to reply personally. This invention makes use of the ability to automatically send a reply to a received message.

**[0041]** The invention may be implemented either in each recipient's computer, or in a central e-mail server, or in an intermediate system, or in any combination of these. Here, just by way of example, it is assumed that the invention is included in the recipient's local system as part of the installed e-mail program, or as a separate filter component that cooperates with this program.

**[0042]** Assume by way of illustration that a sender identified as "John@mail.ext" sends a message to a recipient whose e-mail address is "Fedup@mailfront.com." Assume further that the subject line of the message is:

**[0043]** "Get out of debt!!!\*\*\*\*\*x3zvi1"

Before this original message is passed to and listed in Fedup's e-mail Inbox, a challenge message is automatically generated and sent to John@mail.ext by Fedup's system, in particular, by the active filter according to the invention. The challenge includes a question, or a direction to perform some other action, whose correct answer or action must be part of a response that the sender returns to Fedup. The invention does not pass the original message to Fedup's Inbox unless the sender correctly responds to the challenge.

**[0044]** FIG. 1 illustrates the operation of a simple embodiment of the invention: In response to the original "Get out of debt!!!\*\*\*\*\*x3zvi1" message, the intended recipient's (Fedup's) system returns to "John@mail.ext" a challenge, which, assuming John@mail.ext is even a valid address, would normally be displayed in a window 110 of a display 100 generated by a "John's" conventional e-mail program. (Typical toolbars 112 and 114 are also illustrated.) In this example, the challenge is part of a standard text message, and directs "John" to reply to the question "7x8=?" If, in response, John sends back to Fedup a message with the correct answer 56 as the first line, then Fedup can be almost certain that the sender at John@mail.ext is at least human, as opposed to being a mass-mailing program. The address "John@mail.ext" may then be added to a list of verified, approved, "welcome" senders so that future messages from that address will be accepted with no need for response to another challenge.

**[0045]** The invention may include not only the substantive response criterion that the sender must respond correctly to the challenge, but also one or more formal criteria relating, for example, to the timing and manner of response. Formal response criteria include response within a set response period and/or submission of no more than a maximum number of attempted but incorrect responses. Other formal criteria such as the sending server being configured to accept e-mail and not having open relays may also be chosen and implemented using normal design methods and heuristics.

**[0046]** If "John@mail.ext" responds incorrectly, or fails to meet any other included formal response criterion, then the invention infers that "John" is not a valid or welcome sender, and the original message is either blocked altogether or is placed in a "Reject" folder that Fedup can view if he so chooses. The address "John@mail.ext" is then preferably

also added to list of blocked or “unwelcome” senders, so that future messages from that sender will not be challenged again, but will also not be passed to the recipient.

**[0047]** It is not necessary according to the invention for only complete addresses to be analyzed. Rather, address approval and blocking may be done on the basis of the domain (such as “mail.ext”), regardless of the sender’s name. This is equivalent to allowing “wildcards” in the addresses. Thus, blocking or approving “\*@mail.ext” would correspond to blocking or approving all mail from mail.ext. So as not to block (or pass) all messages from widely used domains, various heuristics should be implemented; these may be chosen using normal design considerations.

**[0048]** The active filtering mechanism according to the invention may be installed at different levels in a recipient’s e-mail “stream.” For example, the active filter may be installed in the central e-mail sever that first receives e-mail over a network from the sender. In this case, the invention may be used to help all users connected to the central server. The invention may also be integrated into each recipient’s local e-mail program, or installed in any conventional manner (such as from a CD ROM or downloaded via the Internet) as a component that cooperates with this program.

**[0049]** The invention may also be installed in intermediate systems, that is, computers operating between each recipient’s local e-mail application and the central e-mail server. For example, an enterprise such as a company or any other organization that maintains for different users a common connection to the general, external e-mail network (such as the World Wide Web) might prefer to install a single copy of the invention in the system that is responsible for the interface with the external network. This can be considered a “perimeter” installation inasmuch as the invention will then operate at the software and hardware “perimeter” separating the enterprise from the public part of the network, which the enterprise does not control.

**[0050]** As is explained in greater detail below, the invention allows not only junk e-mail to be detected and blocked, but e-mail sent from user-specified and even externally compiled addresses and domains may also be blocked. The “perimeter” installation of the invention may then be used to enable the administrator of the enterprise system to better control the flow of e-mail into and out of the enterprise as a whole. Perimeter installation will also reduce the bandwidth drain on the enterprise’s e-mail system even on the way to the central e-mail servers.

**[0051]** The invention may also be included in a suite of junk e-mail filters. Using conventional input methods, such as moving a slider or checking a check box, the user can then adjust the level of security of his e-mail system. A relatively low setting could indicate that only conventional subject-line analysis should be used to filter incoming e-mail. The active filter according to the invention, with challenge-response, would then be the highest setting, since it has the most stringent requirement (accurate, human response) to avoid rejection.

**[0052]** FIG. 2 is a flowchart that shows the main steps performed by the invention: When an e-mail message arrives at the recipient’s system (or at whichever system the invention is installed in), the sender’s address is first compared with the addresses already in the Approved or Blocked lists **360, 362**. If the message is from an approved sender, then the message is passed directly to the user’s e-mail program (for example, Inbox) for display and review. If, however, the message is

from a blocked address, then it may either be deleted or, preferably, placed in a Reject folder. The user may, optionally, also be notified that a message has been blocked so that he may review either the address or the message or both.

**[0053]** If the sender’s address has not been analyzed before, then the message and its address are held in a “Pending” folder or list and a challenge is generated. The challenge may be “standardized,” that is, the same for all senders, until updated by the user, or it may be generated anew for each sender. The example challenge (simple multiplication) in FIG. 1 could, for example, easily be changed for each message to include different multiplicands.

**[0054]** Challenge generation may be automatic, or according to user input, or a combination of both. To continue with the example of a simple two-operand arithmetical challenge such as “7×8=?”, the active filter according to the invention (in particular, a challenge generation module—see below) could generate such challenges from a basic generalization, such as “M Op N=?” where M and N are chosen randomly from some predetermined set of whole numbers and Op is an operator chosen randomly from a group such as addition, subtraction and multiplication. Such randomization would enable the system to easily generate different challenges for different potential junk e-mailers.

**[0055]** The challenge is then sent automatically as a reply to the sender (here, John@mail.ext). The system then waits for the sender’s response to the challenge. A maximum response time is preferably implemented; failure to respond correctly in the allotted time leads to the original message being blocked and the sender’s address being put in the Blocked list.

**[0056]** If the sender does respond, then the response is checked to determine if it is correct. If it is, then the sender’s address is put in the Approved list and the original message is passed to the recipient. It is of course possible that a valid, welcome sender may answer incorrectly by mistake. Moreover, a welcome sender may send the original message and then log off and leave his computer for longer than the expected turn-around time for an immediate response to a challenge. According to the invention, senders may therefore optionally be given more than one opportunity to respond to a challenge before they are blocked and/or they may be given a time for response long enough to take into account normal e-mail habits. If the number of permitted response attempts (which may be one) has not yet been exceeded, and/or the response period has not yet expired, then a new challenge is preferably sent to the sender. If the number of attempts and/or the allowed response time has been exceeded, then the sender’s address is added to the Blocked list and the original message itself is either deleted or placed in the Reject folder.

**[0057]** It would also be possible—but would provide much weaker security—for purely formal criteria to be sufficient. In other words, the challenge “question” might be a “null” question, meaning that any response sent from the same sender address within a set period might suffice to have the sender address be included in the Approved list. For example, the challenge could simply be “Reply to this e-mail message.” This minimum-security challenge could be defeated with an auto-reply feature in the sender’s system, but it would work against the many junk e-mailers who send messages from invalid, or frequently changing, addresses. Even if the sender’s system generates an automatic response, the invention would serve to confirm that the junk e-mail sender’s address is valid and relatively stable—this in itself would make it

easier to identify the sender and possibly stop his junk e-mail activities using other techniques, including legal means.

**[0058]** Even at this point, the advantages of the invention should be clear: Few senders of mass junk mail (especially where the senders are automated) will be able to respond to individual challenges. Even if the junk e-mailer is determined enough to try, responding individually to even a hundred intended recipients (to say nothing of thousands or millions) would greatly increase the burden and cost of sending junk e-mail. As an added advantage, assume that the junk e-mailer sends 10,000 junk e-mail messages from his own computer, using a valid "From" address, and that all of the 10,000 recipients are wise enough to have installed the invention. The junk e-mailer's computer would then receive 10,000 challenge messages, which, one hopes, would inconvenience the junk e-mailer enough to act as a deterrent.

**[0059]** General System Configuration

**[0060]** FIG. 3 shows the main components of the active filter 300 according to the invention, which is included within or cooperates with any conventional e-mail application 310 such as Microsoft Outlook, Eudora, etc., or the conventional e-mail software of any conventional e-mail server. As is mentioned above, the active filter 300 may be included as part of the user's local e-mail application 310, or within the e-mail server 312 via which the user receives and sends e-mail. The active filter 300 is shown as a separate sub-system merely for the sake of clarity. Similarly, at least the e-mail application 310, and preferably the active filter 300 as well, are installed as software components, that is, stored, as executable computer instructions and related data in the user's local computer system 320. Again, they are shown separately for the sake of clarity.

**[0061]** A sender 330 sends an e-mail message to the recipient's e-mail application 310 via a network 340 such as the Internet and the recipient's e-mail server 312. Before the application 310 displays the message, however, it is first passed to the active filter 300. An address analysis module 350 checks the lists of approved, blocked and pending addresses 360, 362, 364, respectively, to determine whether analysis is needed at all. Messages from approved addresses may be passed directly to the user's Inbox (or other designated folder) using the normal routines in the e-mail application 310. Messages from blocked addresses may be sent directly to a Rejected folder 366 (preferably within the e-mail application itself 310).

**[0062]** Any message from an address that is in the Pending buffer 364 is passed to a response analysis module 370, which compares the response (if any) in the message with the correct response. If the response is correct, then the address record is removed from the Pending buffer 364 and may be entered into the Approved list 360. The original message (held in the Pending buffer 364) can then be passed to the e-mail application 310 for normal display to the user. Future e-mail from any address in the Approved list 360 can then be accepted and passed on to the user without further challenges—welcome senders of e-mail will therefore need to correctly answer a challenge only once. If the response is incorrect, the sender may optionally be given yet another try. If a "retry" possibility is implemented, the decision as to whether to allow another try may be based on whether a set response period has expired (which can be determined by comparing the original transmission time against the time and/or date information available from the operating system) and/or whether a maximum number of failed attempts has been exceeded (which can be

tracked by incrementing a counter in the Pending buffer record), or on other formal response criterion.

**[0063]** Note that the invention may be used together with other e-mail filters. Any message that contains a prohibited keyword, for example, or fails some other test, such as having a content or structure indicative of junk e-mail (such as strings of random characters after several blank spaces or asterisks in the subject line, or in the message text), could be automatically rejected and the sender's address may be blocked, regardless of the response.

**[0064]** Assuming that the response is incorrect and that the sender is to be allowed another try, a challenge generation module 372 either resends the original challenge or generates a new challenge that is sent to the sender.

**[0065]** A sender may send more than one message during a given response period. There might therefore be more than one pending record for the same address in the buffer 364. If the sender correctly responds to the first challenge but allows the period for the second to expire, then his address might end up in both the Approved and the Blocked lists. One way for the address analysis module 350 to handle this possibility is to group all pending messages by address, or to scan the Pending buffer for duplicate addresses, whenever a new message is received. The results of the first response analysis corresponding to the duplicated address can then be applied to all messages from the same group—one correct response from a sender will cause all the messages from that address to be "released" for display and response failure will cause all the messages to be rejected.

**[0066]** Another way to handle this possible contradiction is for the address analysis module to periodically compare the Approved and Blocked address lists—any address that is both blocked and approved should probably be approved because the sender will have correctly responded to at least one challenge. Alternatively, the contradictory status can be flagged for the user, who may then manually decide whether to approve or block the ambiguous address.

**[0067]** If the address of the message is new, that is, is not found in any of the lists 360, 362, 364, then the challenge generation module 372 creates and transmits the challenge back to the sender.

**[0068]** The invention preferably includes a user update module 374 with which the user communicates using any conventional input device. Upon user activation (for example by clicking on a displayed icon or selecting this feature from a menu), this module 374 allows the user to view, edit, add and delete addresses in the Approved and Blocked lists 360, 362, and preferably to easily move an address from one list to the other. Moving addresses and messages between e-mail folders using simple "drag-and-drop" is a known feature of most modern e-mail programs. For example, a user's friend may have failed to respond correctly, or on time. He will then notice that his mail never seems to reach the user. In this case the user can simply move the friend's address to the Approved list. Conversely, a junk e-mail system, aware of the invention, may have been lucky enough to guess a correct response; upon seeing the junk e-mail, the user can then manually put the unwelcome address in the Blocked list 362.

**[0069]** Many users want to receive some mass-mailed e-mail, for example, a daily summary of news headlines, or mailings from organizations to which they belong, or even solicitations such as notices of last-minute deals from airlines. Such welcome senders would not be able to respond correctly to a challenge. The user can then enter these known,

welcome addresses directly into the Approved list via the user update module 374 using conventional input techniques.

[0070] FIG. 3 also shows a “Junk E-Mail Database Service” 390 connected to the network 340 and accessible to the user’s e-mail program 310 and therefore to the active filter 300. This optional system may provide different services depending on any given implementation of the invention. In one configuration, the service 390 itself hosts the active filter 300, and filters e-mail on behalf of users who receive their e-mail via the service 390. In another configuration, the service 390 makes available to users updates of the active filter, including challenge updates. These features are described further below.

[0071] FIG. 4 shows an example of the Pending buffer 364, which holds data concerning messages being held pending response (or lack thereof) to a challenge. For each potential junk e-mail message, a record is preferably included in the buffer that has, as fields, data such as the senders’ addresses (addr\_1, . . . , addr\_m), the subject lines and text bodies, the respective challenge(s), correct response(s) (assuming individualized challenges), and statistics such as the time the challenge was sent (the start of the response period), the number of response attempts, etc.

[0072] Challenges

[0073] The challenge illustrated in FIG. 1 is entirely textual, and can be implemented easily in all modern e-mail programs: Regardless of the e-mail formatting used, the challenge will still be understandable, and a simple “Reply” message will suffice to return the response. This challenge format has the added advantage that the user can easily create challenges. The line connecting the user system 320 and the challenge generation module 372 in FIG. 1 illustrates the (optional) ability of the user to select or change challenges. Accordingly, in this manual-challenge embodiment, the user can create new challenges or edit existing ones. For example, the user could change the multiplicands in the challenge “ $7 \times 8 = ?$ ”; he should then enter the correct answer in the response analysis module 370 using any conventional input arrangement.

[0074] Once junk e-mailers become aware of a new filter, they usually set about developing automated ways to defeat it. It would thus be possible for particularly clever junk e-mailers to try to use a parsing program to parse the challenge and automatically produce a correct response. Allowing the user to create challenges and input them (using any conventional input technique) to the challenge generation module 372 may help overcome this potential weakness. If the text challenge requires specific or even general knowledge not commonly within the realm of computers, then automatic response generation will be difficult if not impossible. For example, no parser would correctly respond to the challenge: “What is my daughter’s first name?” Indeed, this is also an example of a challenge that would exclude most human junk-mailers. Other challenges could easily be written to block e-mail from people who do not readily know, say, algebra or Swedish or classical guitar, or a password.

[0075] Most modern e-mail programs allow messages to be sent in or include elements created using mark-up languages such as HTML and image formats such as JPEG. The preferred embodiment of the invention takes advantage of this capability to generate challenges that cannot be parsed at all, because the challenges require correct interpretation of an image. FIG. 5 illustrates an example of such a challenge.

[0076] In FIG. 5, the challenge has been sent to the sender as before, but the challenge message includes an image 500, which includes a challenge question as part of the image. In the illustrated example, two flowers are shown, and the challenge is to correctly indicate the number of flowers. To make responding easy and relatively fail-safe, in the illustrated example, four input fields (the “button” icons labeled 1-4) are included in the image. To respond to the challenge, the sender may simply position a cursor 510 over the button indicating the correct answer and then “click” on the button. By including conventional scripting along with the challenge message, the sender’s (John@mail.ext’s) e-mail program can then automatically send the response back to the recipient (Fedup@mailfront.com) as soon as the sender clicks on his choice of answer. Note that an image-based challenge tests that the sender is human in two ways: First, the sender must see and interpret the picture and the question; second, he must perform the physical acts of positioning the cursor 510 over the correct answer button and then also click a mouse or similar input device.

[0077] One other advantage of image-based challenges (with text also in the image format) that can be answered by clicking on multiple-choice buttons is that they can be used readily for languages whose alphabet, symbols or characters are not supported by standard e-mail programs or browsers.

[0078] FIG. 6 illustrates a challenge that combines image interpretation and a manual, text response. In this case, to respond correctly, the sender must physically type “right” into a response message. FIG. 6 also illustrates how any text included in a challenge image is preferably put in a font, color, style, etc., and/or is “obscured” with other graphics (such as the surrounding dashed boxes) so as to defeat any attempts to “decipher” the image using optical character recognition.

[0079] Other non-textual challenges may also be generated according to the invention. For example, the challenge could include an audio file (for example, in MP3 format), or a moving or changing image (in, for example, MPEG format), with a challenge question or action that presupposes an ability to interpret and/or react to the audio or active-image challenge. The challenge question could, for example, be spoken in the audio file, or could be presented as scrolling text. These non-textual and indeed non-visual challenges may of course be combined with the text- and image-based challenges described above.

[0080] In the preferred embodiment of the invention, the challenge generation module 372 automatically chooses which challenge to issue to senders, for example, from a database of challenge texts and/or images, etc., included as part of the active filter. One advantage of this is convenience: The user does not need to think of appropriate challenges, so that different challenges can easily be issued to even a large number of different sender addresses. Note that it will often not be necessary to have available a large number of different challenges—even a single “standard” challenge will still require human interpretation and action for correct response. Whenever a new challenge is required, the module 372 may then, for example, choose it randomly from the database. Another advantage is that automatic generation ensures that challenges will be in the correct format. Still another advantage is that new challenges can be downloaded from a provider via the Internet, for example as part of a periodic update routine. Such automatic updating via the Internet is a well known procedure used, for example, by all major providers of anti-virus software.

**[0081]** Much junk e-mail is sent from forged addresses. Any challenge sent to such an address will typically result in the return of a “bounce” error message, that is, an e-mail message stating that the challenge was undeliverable. The From address of such an error message is usually “postmaster@domain.ext” or “route@domain.ext” where “domain.ext” is the domain name and extension used in the junk e-mail message. The text of the undelivered message is, moreover, typically included as part of the error message, as well as the address to which the message could not be delivered. This provides yet another way for the response analysis module **370** to determine that a message is junk: Whenever a message is received from, for example, “postmaster” or “route,” then the module can check the domain name and text to determine which challenge the error message corresponds to. The sender address that caused the bounce can then be assumed to have sent a junk e-mail and can therefore be placed in the Blocked list.

**[0082]** The response analysis module **370** preferably also immediately blocks the sender address and the original e-mail message if, instead of a correct response to the challenge, an auto-reply message is returned. That a message is generated by auto-reply can be determined using known methods. For example, according to current e-mail standards, all auto-reply messages have an “x-” header that specifies that the message is “precedent bulk.”

**[0083]** Still another situation that the active filter preferably automatically interprets as indicating junk e-mail is when an incoming e-mail message has, as its “From” address, the user’s own address. In other words, if the “From” address is identical to the “To” address, then the message can be assumed to be junk. Because blocked messages and their addresses are preferably stored in the Reject folder and Blocked list for later review, the user will still be able to retrieve any message that he deliberately sent to himself, for example as the result of sending a “Reply to all” message. Alternatively, by keeping track of identifying information for any “Reply to all” message sent, for example, the message body, the active filter could automatically approve without challenge any received message that has the identifying information.

**[0084]** Anti-“Ping-Ponging”

**[0085]** Assume that the invention is installed in the computers of two different users U1, U2, neither of whom are junk e-mailers. When U1 sends a first message to U2, U2’s computer will return a challenge C1 to U1. U1’s active filter will interpret this challenge C1 either as an incorrect response, or as incoming e-mail, which will cause it to issue a challenge C2 to U2, whose active filter will interpret C2 either as an incorrect response to C1, causing blocking of U1’s address, or as yet another message, causing generation of yet another challenge C3, and so on. The result will therefore be either that U1 and U2 block each other’s addresses, or that their active filters will begin an endless series of “ping-ponging” challenges.

**[0086]** One way to prevent, or at least cut off, ping-ponging is for the active filter to set a limit on the maximum number of challenges issued to any given address during some predetermined time period. If the limit is reached, and if all messages received from the sender are at least substantially identical, and if they are not auto-replies, then the active filter can assume that the two systems are ping-ponging challenges to each other. The active filter can then handle this situation in different ways depending on user or designer preferences. For

example, the first returned message in a ping-pong sequence can then be passed to the Inbox for review. Alternatively, if the active filter determines that ping-ponging is happening, then after a predetermined number of “rounds,” the active filter could notify the user, for example, using a pop-up window, indicate the remote address involved, and prompt the user to decide whether to accept the original message or reject it (and block the sender’s address) without further challenge.

**[0087]** In the preferred embodiment of the invention, ping-ponging is prevented by including a key in each challenge generated, for example, in the header. Upon receipt of any message that includes a valid key, the active filter will then know that the incoming message is a challenge and may then automatically approve the sender address and pass the message to the user’s Inbox without issuing a challenge.

**[0088]** The key may be a single number generated using, for example, a weak algorithm such as a checksum, or a strong algorithm such as those used to for public-key cryptography, or any other known encryption routine. The basis of the key is preferably all or some portion of the outgoing (challenge) message, including the From and To addresses. Other bases may of course also be used, such as those that are functions of the time, or some random number. This information will be available to the active filter in the receiving system, so that it too can calculate the key. If the key is correct, then the recipient system knows that the message was sent by a system that incorporates the invention and can probably be trusted.

**[0089]** It would be possible for particularly determined and skilled junk e-mailers to install the invention in their own systems, then generate false messages to themselves “from” the addresses to which they wish to send junk e-mail. The “challenge” their systems would then send would be the junk mail message, with a valid key, which might pass through the active filter of the recipient’s system at the “real” addresses. The recipient’s system will not have sent any message to the “From” address of the junk e-mail, however, so this fact can be used to override the key and block the message and junk e-mailer’s address: Since the key is preferably included only in issued challenges, the recipient should not receive any “challenge” with a valid key from any address to which it has not initiated e-mail exchange. An incoming message (usually a challenge) can therefore be passed directly to the Inbox of the recipient if the incoming message has a valid key in its header and the recipient’s system earlier initiated a message to the sender of the incoming message.

**[0090]** Automatic generation of challenges also helps prevent this deception, since it can ensure that all challenges must conform to a predetermined format. Even if a correctly calculated key is found in a message header, but the message body does not conform to the predetermined format, then the active filter will know that the incoming message is not a valid challenge. Other conventional filter mechanisms may also be employed in this case to detect junk content.

**[0091]** Imported and Dynamic Address Lists

**[0092]** In the embodiments of the invention described above, a sender’s address is included in the Approved list **360** when the sender correctly responds to a challenge. This is not the only way to build up this list in the preferred embodiment of the invention. To avoid sending challenges to senders who are already known to be welcome, the active filter **300** according to the invention preferably puts in the Approved list **360** any addresses to which the user himself manually sends or has sent original (and, possibly, reply) e-mail messages. The filter

**300** could also import the addresses in the “address book” of the user’s e-mail program **310**. The Approved list can thus be built up dynamically.

**[0093]** In another embodiment of the invention, externally compiled, approved addresses are imported for inclusion in the user’s list **360**. Assume for example that the user is a new employee at a company. His computer’s Approved list could then be loaded initially with the addresses of all the other employees, of known customers and vendors, etc. Such company-wide Approved lists could also be updated periodically by a system administrator. As the user sends and replies to e-mail, his Approved list will therefore grow dynamically to include addresses approved both “globally,” that is, from the company-wide list, and “locally,” that is, approved as the result of correct challenge responses.

**[0094]** Addresses known to be globally unwelcome or prohibited could similarly be downloaded or stored in the Blocked list **362**. These “globally” blocked addresses could also be made “undeletable” by marking them appropriately and coding the user update module **374** so that it does not allow the user (but rather only an administrator) to remove these addresses from the Blocked list **362** or move them to the Approved list **360**.

**[0095]** Dynamic compilation of the Blocked list **362** may also be used in implementations in which the active filter **300** is included in the central e-mail server **312**, to which many users are typically connected. If the address of a junk e-mailer is blocked centrally, not only would this prevent the junk e-mail from reaching all those connected to the server, but it would also prevent the waste of bandwidth. The users connected to the server could also be allowed to participate in building up the Blocked list **362**. In this case, for example, if more than a minimum number of users indicate to the central sever that a particular address is used for junk mail, then the central server could include it in the global Blocked list. In order to prevent a coordinated “attack” against a legitimate sender, the administrator of the central server should preferably review all such globally blocked addresses.

**[0096]** FIG. 3 also illustrates an extension of the concept of a central list of blocked addresses. As shown in FIG. 3, a third-party “Junk E-Mail Database Service” **390** may be included and made accessible to the user via the network **340**. According to this embodiment of the invention, the service **390** preferably includes a copy of the active filter **300**, as well as any other conventional junk e-mail filters. The service preferably not only detects junk e-mailers’ addresses (and/or domains) using the invention, but also actively seeks them out. Regardless of the methods used to identify junk e-mailers’ addresses, the service **390** includes a Blocked list, preferably in the same format as those used in the active filter **300** of user systems. The Blocked list in the service **390** is therefore a global list that may be accessed by users and downloaded into their respective Blocked lists **362**. The service **390** may provide blocked address lists by subscription, or it may be associated with the e-mail service provider (usually, but not necessarily, the same as the user’s Internet Service Provider). Users who subscribe to the service **390** may then download the updated global Blocked list, which is added to the locally blocked addresses in each user’s respective system.

**[0097]** FIG. 7 illustrates one example of a graphical interface that allows the user to control the active filter according to the invention, and that also allows for prompted challenge generation. In the illustrated example, the user is writing an

e-mail message to Brian@mailfront.com using a standard e-mail program; the interface may be included in other applications as well, for example, a browser. A window **710** is shown on a display screen **700**, along with toolbars **712** and **714**. In this example, the user accesses the features of the active filter via a drop-down menu **720**, which is activated by clicking on a “Verify” tool on toolbar **712**. Any other conventional access arrangement may of course be implemented instead.

**[0098]** The drop-down menu allows the user to turn the active filter on or off, that is, to enable or disable the challenge-response mechanism of the invention, and, if implemented, to activate prompted challenges. Additional features that the user may be allowed to access include the Approved and Blocked lists so that addresses stored in these lists may be added, deleted, edited or moved.

**[0099]** When prompted challenging is activated, the active filter does not automatically generate challenges and analyze responses for e-mail messages from as yet unclassified (approved, blocked or pending) sender addresses. Instead, when the user’s e-mail program receives a message from an unclassified address, the challenge generation module **372** (or some other dedicated module) causes a pop-up window **730**, call-out, etc. to be displayed to notify the user that unclassified e-mail is incoming, that is, has been received but not yet approved for display. The address of the incoming e-mail and, optionally, all or part of the subject line, are also preferably displayed. By clicking on the desired choice, the user can then select to directly approve the address and permit the message to be entered in his Inbox, to block the address and reject the message, or to instruct the active filter **400** to generate a challenge and act according to any response or lack of response. In FIG. 7, for example, the user has chosen to have the message from John@mail.ext challenged.

What is claimed is:

1. A method for filtering electronic mail (e-mail) sent to a user comprising:
  - receiving a current e-mail message from a sender address; via e-mail, returning to the sender address a challenge that requires a correct e-mail response; and
  - blocking the current e-mail message from the sender address unless the correct e-mail response is received by the user system from the sender address.
2. A method as in claim 1, further including the step of blocking future e-mail messages from the sender address unless the correct e-mail response is received by the user from the sender address.
3. A method as in claim 1, in which the challenge includes an instruction to perform an action, the correct e-mail response requiring correct performance of the action.
4. A method as in claim 3 including the sub-step of including a question in the challenge, in which the action is answering the question.
5. A method as in claim 3 further comprising the sub-step of including an image in the challenge.
6. A method as in claim 5 including the sub-step of incorporating a legible question as part of the image, in which the action is answering the question.
7. A method as in claim 5 further comprising the sub-step of including in the challenge, in a text format, a question about the image, in which the action is answering the question.
8. A method as in claim 3, further comprising user selection of the challenge.

9. A method as in claim 3 further comprising the sub-step of including an audio file in the challenge.

10. A method as in claim 3 further comprising downloading challenges via a network from an external challenge-generation system and selecting for sending to the sender address at least one of the downloaded challenges.

11. A method as in claim 3, further including the step of blocking the current e-mail message from the sender address unless the correct e-mail response is received according to a formal response criterion.

12. A method as in claim 11, in which the formal response criterion is receipt by the user from the sender address of an e-mail response within a response period.

13. A method as in claim 11, further comprising allowing multiple attempts to return a correct e-mail response from the sender address by submitting at least one additional challenge to the sender address upon receipt of an incorrect e-mail response from the sender address.

14. A method as in claim 13, in which the formal response criterion is receipt by the user from the sender address of the correct e-mail response within at most a predetermined maximum number of the multiple attempts.

15. A method as in claim 3 further comprising the following steps:

upon receipt of the correct e-mail response from the sender address, including the sender address in a list of approved sender addresses and accepting the current e-mail message; and

accepting without challenge any e-mail message sent from an address included in the list of approved sender addresses.

16. A method as in claim 15, further comprising the step of including user-specified e-mail addresses in the list of approved sender addresses.

17. A method as in claim 15, further comprising the step of including in the list of approved sender addresses any e-mail address to which the user has initiated sending an e-mail message.

18. A method as in claim 15, further comprising the step of including in the list of approved sender addresses e-mail addresses obtained from an externally compiled address list.

19. A method as in claim 3 further comprising the following steps:

unless the correct e-mail response is received from the sender address, including the sender address in a list of blocked sender addresses; and

blocking without further challenge any e-mail message received from any address included in the list of blocked sender addresses.

20. A method as in claim 19, further comprising the step of including user-specified e-mail addresses in the list of blocked sender addresses.

21. A method as in claim 19, further comprising the step of including e-mail addresses obtained from an externally compiled address list in the list of blocked sender addresses.

22. A method as in claim 19, further comprising the step of including the sender address in the list of blocked sender addresses unless the correct e-mail response is received according to a formal response criterion.

23. A method as in claim 19, further comprising the step of providing the list of blocked addresses to third parties.

24. A method as in claim 3 further comprising the following steps:

notifying the user of receipt of the current e-mail message and prompting to indicate whether to send the challenge; and

sending the challenge only if the user indicates that the challenge should be sent.

25. A method as in claim 1 further including the steps of calculating, for each challenge, a corresponding key, and including the key in the challenge.

26. A method as in claim 25 further including the step of accepting for display to the user, without challenge, any incoming message that contains a correctly calculated key.

27. A method as in claim 25 further including the step of accepting for display to the user, without challenge, any incoming message that both contains a correctly calculated key and is sent from a sender address to which the user previously initiated sending a message.

28. A method for filtering electronic mail (e-mail) sent to a user comprising:

receiving a current e-mail message from a current sender address;

via e-mail, returning to the current sender address a challenge that requires a correct e-mail response, the challenge including an instruction to perform an action and the correct e-mail response requiring correct performance of the action;

if the correct e-mail response is received from the current sender address, including the current sender address in a list of approved sender addresses and accepting without challenge the current e-mail message and any e-mail message sent from an address included in the list of approved sender addresses; and

if the correct e-mail response is not received from the current sender address, including the current sender address in a list of blocked sender addresses and blocking without further challenge any e-mail message received from any address included in the list of blocked sender addresses, including the current e-mail message.

29. A system for filtering electronic mail (e-mail) sent to a user comprising:

a computer that is connected to a network over which e-mail is transmitted;

a computer program executing on the computer, wherein the computer program comprises computer instructions for:

receiving a current e-mail message from a sender address to a user address;

via e-mail, returning to the sender address a challenge that requires a correct e-mail response; and

blocking the current e-mail message from the sender address unless the correct e-mail response is received by the computer from the sender address.

30. A system as in claim 29, in which the computer program further comprises computer instructions for generating the challenge to include an instruction to perform an action, the correct e-mail response requiring correct performance of the action.

31. A system as in claim 30, further comprising an approved list storing approved sender addresses, in which the computer program further comprises computer instructions for:

including the sender address in a list of approved sender addresses and for accepting the current e-mail message



upon receipt of the correct e-mail response from the sender address; and

accepting without challenge any e-mail message sent from an address included in the list of approved sender addresses.

**32.** A system as in claim **30**, further comprising a blocked list storing disapproved sender addresses, in which the computer program further comprises computer instructions for:

including the sender address in a list of blocked sender addresses unless the correct e-mail response is received from the sender; and

blocking without further challenge any e-mail message received from any address included in the list of blocked sender addresses.

**33.** A system as in claim **29**, further comprising a challenge generation module generating the challenge to include an instruction to perform an action, the correct e-mail response requiring correct performance of the action.

**34.** A system as in claim **33**, further comprising a response analysis module analyzing any e-mail response from the sender address and determining whether the e-mail response is correct.

**35.** A system as in claim **33**, in which the challenge generation module is provided for downloading challenges via the network from an external challenge-generation system and selecting for sending to the sender address at least one of the downloaded challenges.

**1.** A method for filtering electronic mail (e-mail) sent to a user comprising:

receiving a current e-mail message from a sender address; via e-mail, returning to the sender address a challenge that requires a correct e-mail response; and blocking the current e-mail message from the sender address unless the correct e-mail response is received by the user system from the sender address.

**2.-35.** (canceled)

\* \* \* \* \*