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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 April 2010 (01.04.2010)

PCT

(10) International Publication Number
WO 2010/036612 A1

(51) International Patent Classification:
H04L 29/06 (2006.01) *H04L 29/08* (2006.01)

(21) International Application Number:
PCT/US2009/057685

(22) International Filing Date:
21 September 2009 (21.09.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/100,192 25 September 2008 (25.09.2008) US
12/537,760 7 August 2009 (07.08.2009) US

(71) Applicant *(for all designated States except US)*: **NTT DOCOMO, INC.** [JP/JP]; Sanno Park Tower, 11-1, Nagata-cho, 2-chome, Chiyoda-ku (JP).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KOZAT, Ulas, C.** [TR/US]; 3612 Flora Vista Avenue, #349, Santa Clara, CA 95051 (US). **RADOSAVAC, Svetlana** [RS/US]; 655 South Fair Oaks Avenue, #a-305, Sunnyvale, CA 94086 (US). **KEMPF, James** [US/US]; 369 Foxborough Drive, Moutain View, CA 94041 (US).

(74) Agents: **BERNADICOU, Michael, A.** et al.; Blakely, Sokoloff, Taylor & Zafman LLP, 1279 Oakmead Parkway, Sunnyvale, CA 94085-4040 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: A METHOD AND APPARATUS FOR SECURITY-RISK BASED ADMISSION CONTROL

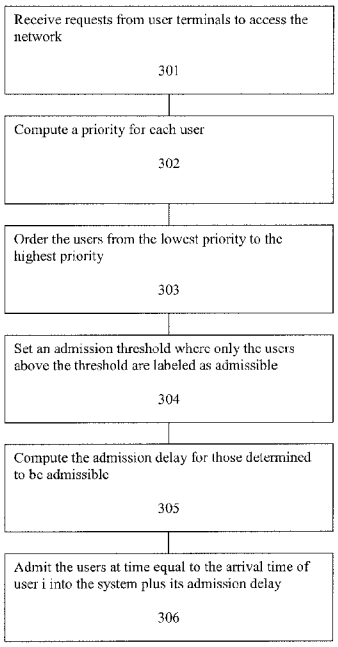

Figure 3

(57) Abstract: A method and apparatus is disclosed herein for security risk-based admission control. In one embodiment, the method comprises: receiving a request from the user device to access the network; determining whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, wherein the constraint optimization is based on a utility associated with admitting the user device to the network, a reputation value associated with the user device, and a botnet damage estimation on the network associated with the user device; and admitting the user device to the network based on results of determining whether to admit the user device.

WO 2010/036612 A1

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

A METHOD AND APPARATUS FOR SECURITY-RISK BASED ADMISSION CONTROL

PRIORITY

[0001] The present patent application claims priority to and incorporates by reference the corresponding provisional patent application serial no. 61/100,192, titled, "A Method and Apparatus for Security-Risk Based Admission Control", filed on September 25, 2008.

FIELD OF THE INVENTION

[0002] The present invention relates to the fields of communication networks and network security; more particularly, the present invention relates to admission control that decides to admit user devices into a network based on security risk assessments of user devices waiting to connect the network and those already in the network.

BACKGROUND OF THE INVENTION

[0003] Admission control in a communication network amounts to the decision of admitting a particular subscriber/user into the system or not every time that subscriber/user wants to connect to the network and use network's resources. The traditional admission control policies typically address authorization, authentication, and quality of service issues in their decision processes.

[0004] Security risks arise due to several facts such as: subscriber's private information can be compromised, subscriber devices are compromised and are used to launch attack on others by spreading malware, many subscriber devices together form a botnet to launch distributed denial of service attacks on the network itself and/or on other network users, mainly businesses, etc. The damages can be measured both in monetary and network performance terms. The network performance can be measured as the disconnectivity incurred inside the network or at the edge links of the network that connects businesses and users to the communication network, e.g., Internet.

[0005] Cyber insurance was first proposed as a method for mitigating the residual risk in the Internet in 2000. The cyber insurance policy was offered through a partnership of two companies: security company Counterpane and insurance company Lloyd's of London.

[0006] There are a number of problems that arise in this field, which are analogous with the auto and health insurance markets. Others have stated that, just like other successful insurance markets, the cyber insurance market will be developed over time in a response to experience and result in well-functioning insurance markets.

[0007] An analysis of the impact of insurance and self-investment in user-user interactions has been developed. This analysis indicates that protection against attacks involves four different responses: 1) avoid the risk, 2) absorb the risk, 3) self-protect to mitigate the risk, and 4) transfer the risk through insurance or hedging. The analysis starts with a utility function model of the interplay between insurance and self protection in the single agent case. Depending on the costs of self protection and insurance versus the probability and perceived loss from attack, users are motivated to either insure and seek self-protection, not insure and seek self-protection, or absorb the risk by not protecting. The analysis then extends this model to multiple agents and examines the effects of moral hazard (the tendency of people to engage in more risky actions when they believe their losses will be compensated). This multi-agent model also considers the interactions between self-protection in different individuals, where the decision of one individual to self-protect affects the losses of others in the case of attack. The analysis applies this analysis to two different kinds of networks, a full mesh and a star network very similar to the Internet's power law network form. They observe a threshold phenomenon, in which the reduced premiums for self-protection for insured users cause a small portion of the population to invest in self-protection, which ultimately causes all users to self-protect.

[0008] Others have assumed that the security risk of each player in the network depends on a linear combination of investments of all users in the network and have shown that the Price of Anarchy (POA) is very large in the one-shot game and increases with the number of players. In the repeated game, it is possible to

achieve social optimum if it doesn't interfere with individual rationality. However, implementing this strategy in a repeated game requires cooperation and communication among the players. This can be achieved either in an environment where all players cooperate or when a social planner that ensures certain level of investments by all users is used.

[0009] The applicability of the existing insurance schemes in the current Internet has been explored. Using standard insurance models, whether business models based on cyber insurance schemes that utilize such insurance models can survive in the competitive insurance market have been analyzed. After taking into account information asymmetry (before contract signing) and hidden information (after contract signing), it would appear that no policy that is based on the current insurance models can survive in the competitive market. Therefore, different Internet architecture must be adopted for mitigating and/or eliminating this information asymmetry.

SUMMARY OF THE INVENTION

[0010] A method and apparatus is disclosed herein for security risk-based admission control. In one embodiment, the method comprises: receiving a request from the user device to access the network; determining whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, wherein the constraint optimization is based on a utility associated with admitting the user device to the network, a reputation value associated with the user device, and a botnet damage estimation on the network associated with the user device; and admitting the user device to the network based on results of determining whether to admit the user device.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] 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.

Figure 1 illustrates one embodiment of a network with user devices and an admission controller.

Figure 2 is a flow diagram of one embodiment of a process for admitting a user device waiting to enter a network having one or more currently admitted user devices, based on a security assessment of the user devices.

Figure 3 is a flow diagram of one embodiment of a process for admitting one or more user devices into a network.

Figure 4 is a block diagram of an exemplary computer system.

Figure 5 is a flow diagram of one of embodiment of the admission control process.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

[0012] An admission control policy for communication and computer networks is disclosed. In one embodiment, the admission policy is targeted towards building a secure network, where the security is not rigid in the sense that network and users can bear a certain degree of security risk, where the cost of feasible attacks is marginalized and/or compensated (e.g., through insurance).

[0013] In the following description, numerous details are set forth to provide a more thorough explanation of the present invention. 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.

[0014] 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.

[0015] 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.

[0016] 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.

[0017] 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.

[0018] 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; and/or flash memory devices.

Overview

[0019] A security-based admission control policy for maintaining secure networks is disclosed. In one embodiment, the admission control policy is used to establish a secure network whose resources such as links, routers, switches, data centers, storage, services, etc. can be used by only the users admitted into the system. The network to be secured might be a private network, it can be a public network with users carrying devices that have installed right set of hardware and software solutions allowed and/or controlled by the network, or it can be a network with multiple isolated virtualized slices running over the same set of physical resources without interfering with each other and some of the slices run the admission control policies as disclosed herein. In one embodiment, the admission policy is targeted towards building a secure network, where the security is not rigid in the sense that network, and users can bear a certain degree of security risk, where the cost of feasible attacks is marginalized and/or compensated (e.g., through insurance).

[0020] In one embodiment, the admission control policy disclosed herein uses the security risks as a system constraint and tries to maximally improve the utility of admission decision. In one embodiment, the admission control policy not only determines whether a user is admissible or not, but when the user is admissible as well.

[0021] Figure 1 illustrates one embodiment of a network with user devices 101-106 and an admission controller 120 which determines whether to admit user devices into the network based on a policy. User devices 101-106 may be mobile or

fixed stations or devices. Although only 6 are shown, the number of user devices may be more or less than this number. Note that for purposes herein, the terms "user device" and "user" will be used interchangeably with respect to their admittance into the network.

[0022] Using the policy, an admission controller 120 admits network users based on their security assessment and imposed security risk on the network and other network users. In one embodiment, admission controller 120 assesses the risk based on the past behavior of the user as well as the real-time scanning and device inspection before the admission into the system. Admission controller 120 manages the accessed network and makes an admission decision for each candidate network user by computing the overall risk after the admission of that particular user. If the overall risk is below a certain threshold that can be accommodated by the network operator, admission controller 120 admits the user; otherwise, admission controller 120 delays the user's admission is delayed until the overall risk drops below the threshold. In one embodiment, the overall risk can be reduced due to (1) increased reputation of the users who are already in the system, (2) departed users, (3) increased reputation of the user waiting to be admitted during the scanning and inspection period, (4) capacity increase inside the network and at the edges, (5) accumulated wealth by the operator due to the premium charged to access the network, and so on.

[0023] When there is more than one user waiting to get admitted at a given time, admission controller 120 decides when to admit each user based on the system constraints and utility of admitting each user at a particular access delay. In one embodiment, admission control 120 uses a security-based admission control policy for maintaining secure networks, where the admission control policy uses the security risks as a system constraint and tries to maximally improve the utility of admission decision. In one embodiment, the utility reflects the service level agreement or quality of service, while the system constraints reflect the quality of security for the admitted users.

[0024] In one embodiment, admission controller 120, using the admission control policy, not only determines whether a user is admissible or not, but when the user is admissible as well.

[0025] The network of Figure 1 also includes monitoring device 121 that monitors network traffic and/or congestion as well as a scanning device 122 that scans user devices seeking entry into the network and downloads security patches (or causes the download of security patches) to these user devices based on the results of their scan. Note that admission controller 120, monitoring device 121, and scanning device 122 may be one device or multiple devices and may comprise a client or server that contains a network interface or some other method that can be used to provide connectivity to the network, and may be located anywhere inside the network.

[0026] Figure 2 is a flow diagram of one embodiment of a process for admitting a user device waiting to enter a network having one or more currently admitted user devices, based on a security assessment of the user devices. The process is performed by processing logic that may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the processing logic is part of or controlled by an admission controller.

[0027] Referring to Figure 2, the process begins by processing logic receiving a request from the user device to access the network (processing block 201). In response to the request, processing logic scans the user device to perform security checks and downloads any necessary software based on scanning results (processing block 202). In one embodiment, where the security assessment is based on a reputation value of a user device, processing logic increases reputation value of the user device responsive to the scanning of the user device and the downloading of software updates.

[0028] After receiving a request to enter a network, processing logic determines whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, where the constraint optimization is based on a utility associated with admitting the user device to the network, a reputation value associated with the user device, and a

botnet damage estimation on the network associated with the user device (processing block 203). In one embodiment, the security assessment is based on a reputation value. In one embodiment, processing logic determines to admit the user device to use the network occurs if results of determining whether to admit the user device occurs indicate a likelihood of possible damage that can be incurred by the user device and the one or more currently admitted user devices to disrupt one or more network services is lower than a threshold. In one embodiment, the likelihood of possible damage comprises the likelihood that the user device and the one or more currently admitted user devices could cooperate together to become a botnet. In one embodiment, determining whether to admit the user device is based on a traffic injection rate of the user device, a reputation value indicative of a trust level for the user device with respect to the network, and each reputation value of the one or more currently admitted user devices.

[0029] Once a determination has been made to admit a user device, processing logic determines when the user device is to be admitted (prior to admitting the user device) (processing block 204).

[0030] Next, processing logic admits the user device to the network based on results of determining whether to admit the user device (processing block 205). Note that processing logic may determine whether to admit the user and admit the user occurs prior to completion of both scanning the user device to perform security checks and downloading software updates to address security risks identified from the security checks.

[0031] Figure 3 is a flow diagram of one embodiment of a process for admitting one or more user devices into a network. The process is performed by processing logic that may comprise hardware (e.g., circuitry, dedicated logic, etc.), software (such as is run on a general purpose computer system or a dedicated machine), or a combination of both. In one embodiment, the processing logic is part of or controlled by an admission controller.

[0032] The process begins by processing logic of the admission controller receiving requests from user devices to access the network (processing block 301).

[0033] Using a utility function, processing logic of the admission controller computes a priority for each user (processing block 302). In one embodiment, the

priority is based on the maximum admission delay and a derivative of the utility function. Other concave utility functions can be similarly defined and for each a different priority function can be computed.

[0034] Next, processing logic of the admission controller orders the users from the lowest priority to the highest (processing block 303). After the priority is computed, processing logic of the admission controller sets an admission threshold where only the users above the threshold are labeled as admissible (processing block 304).

[0035] Processing logic of the admission controller computes the admission delay for those determined to be admissible (processing block 305). Then, processing logic in the admission controller admits the users at time equal to the arrival time of user i into the system plus its admission delay (processing block 306).

Specific Algorithm Embodiments

[0036] In one embodiment, there are K user devices that have already been admitted into the network and N user devices waiting to be admitted. The utility of the network increases as more users are admitted. In one embodiment, the goal of the network is to admit as many users as quickly as possible while keeping expected damages below a tolerable threshold. In order to achieve this, the network constructs an admission policy that admits users based on a security assessment and the security risk they impose to the network and to the already admitted users.

[0037] Each user device u_i, $i \in \{1,\ldots,N\}$ that attempts to join the secure network is characterized with two parameters: reputation p_i, $p_i \in [0, p_{i,\max}]$ and traffic injection rate r_i (i.e. the user is allowed to inject traffic no faster than this rate). While for some cases, $p_{i,max} < 1$, in one embodiment, it is assumed that $p_{i,max} = 1$. Reputation p_i signifies the trust level put on that user by the network. When user u_i requests access to the secure network, the system determines its initial reputation value $p_{i,0}$ based on the results of real-time scanning the user device, past interactions between the user and the network, etc. In another embodiment, if an access control list exist, which shows what privileges user has to access network and networked resources, it can be also used, particularly since the more privileges a user has the more damage it can cause and hence the higher the risk is.

[0038] After $p_{i,0}$ is determined for all user devices waiting for admission, the admission control policy mechanism of the admission controller makes an admission decision for each user device by computing the overall risk after the admission for that particular user. The reputation of an admitted user is assessed and updated in real-time starting from an initial value at the time of the user's arrival:

$$p_i = p_{i0} + g(\tau_i), i = 1,\ldots,N \quad (1)$$

where $g(\tau_i)$ is a non-negative non-decreasing function of admission delay τ_i, i.e., as the access delay increases the system either discovers that the specific user is secure and the reputation increases or discovers that the user does not have the required properties and forces the user to update the device, which results in increase of the user's reputation (otherwise the reputation remains the same).

[0039] The admission controller also captures the security threat of a given subset $\mathbf{B_i} = \{u_{i1}, u_{i2}, \ldots, u_{im}\}$ of admitted user devices u_{i1} to u_{im} by the sum rate $\Sigma_i = \sum_{j=1}^{m} r_{ij}$ where r_{ij} is the rate user u_{ij} can connect to the network, i.e., the user device is allowed to inject traffic at most at this rate. The damage $D(B_i)$ that can be caused by subset B_i is then modeled as a non-decreasing monotonic function of Σ_i, i.e., $D(B_i) = f(\Sigma_i)$. In one embodiment, the form of f is a linear function of Σ_i, but it can take other arbitrary forms as long as it is non-decreasing with the sum rate. In one embodiment, the function f that is used is linear $f(\Sigma_i) = c \times \Sigma_i$ for some constant "c".

[0040] In one embodiment, a reputation value per user (e.g., $0 \leq p_j \leq 1$) that signifies the trust level put on that user by the system is used. The trust is continuously updated based on the past interactions and reputation of the user as well as based on the real-time scanning of the user device. The quantity $(1- \boldsymbol{p_j})$ is then used to measure the likelihood/probability of user **j** becoming a member of a malicious subset $\mathbf{B_i}$. Assuming the misbehavior of one arbitrary user is independent of another arbitrary user, the damage probability $\boldsymbol{\pi_i}$ of a particular subset of users $\mathbf{B_i}$ is computed as

$$\prod_{j \in B_i} (1 - p_j)$$

in one embodiment.

[0041] In one embodiment, the admission control policy guarantees that expected damage over all possible subsets of admitted users is less than a threshold $\mathbf{D_{th}}$, which is computed based on the a-priori assessment of the cost of damage that is tolerable to the network operator and end users. The computation of $\mathbf{D_{th}}$ involves many considerations such as the service level agreement (SLA) between the users of the network and the network operator, premiums paid to access the network, premiums paid to get insurance based on this threshold, the compensation amount by the operator if the SLA is violated and/or an insurance payment is made, the accumulated wealth of the operator and insurer, future cost implications on the victim of the security attack, etc. In one embodiment, the admission controller computes the expected damage as:

$$E_B[D] = \sum_{i \in B} \pi_i \cdot D(B_i),$$

where the expectation is taken over the set **B** of all possible subsets $\mathbf{B_i}$.

[0042] In one embodiment, the admission controller uses the admission control policy to determine which user to admit at what time by solving the following optimization question:

$$\max \sum_{i=1}^{N} U_i(\tau_i)$$

s.t.

$$E_B[D] \le \Gamma_{th} \qquad (1)$$

$$p_i = p_{i0} + g(\tau_i), \quad i=1,\ldots,N \qquad (2)$$

$$0 \le p_i \le 1, \; i=1,\ldots,N \qquad (3)$$

In other words, the objective of the admission controller is to maximize the sum utility of users admitted into the system, where the individual utilities depend on the admission delay of that particular user. As set forth above, *N* denotes the number of users that are not yet admitted into the network. The admission control policy can be executed periodically or after each new user arrival. When a user arrives at time t, there are already other users either already admitted into the system or waiting for their admission time as computed by the solution of the utility optimization presented above. The number of users N in the optimization problem then

corresponds to the total number of users whose arrival time plus the previously computed admission delay exceeds the current time. In other words, N corresponds to the users who are not already admitted into the system and waiting to get admitted. Therefore, at each new arrival, N might be different and for some users admission control algorithm can be executed for more than once according to an updated utility function.

[0043] The first constraint (1) reflects the cost of expected damage over the set of possible attack points on the system, the second constraint (2) reflects the fact that the reputation p_i of each user device is assessed and updated in real-time starting from an initial value at the time of user's arrival, while the last constraint (3) reflects the fact that user reputation is less than 1 and greater than zero. The function $g(\tau_i)$ is a non-negative increasing function of the admission delay τ_i. The second constraint means that p_i is a non-negative non-decreasing function of the admission delay, i.e., it cannot be smaller than the initial reputation, and it increases with time. As soon as a user requests to connect to the system, it starts with an initial reputation value securely obtained by off-line evaluation and the admission controller begins scanning the user device.

[0044] In one embodiment, as time progresses and user passes certain scanning steps as secure, its reputation increases. When the user device passes the scan, its reputation value increases. When a step is not passed, the user device is quarantined and the necessary security patches (after being downloaded) are installed, after which the admission controller increase to the reputation value for the user device because the user device is now up to the current security level. User devices do not need to wait until a full scanning is completed and/or necessary patches are applied. If their current risk imposed on the network as captured by constraint (1) is acceptable and there is greater gain of admitting the user earlier than the full scan in accordance with the utility maximization, then the user can be admitted earlier. However, the scanning process and reputation update continues even after the user is admitted to the system.

[0045] In one embodiment, $g(\tau_i)$ is a linear function of the form $g(\tau_i)=\alpha\times\tau_i$, where α is a positive constant. For a maximum scanning delay target

of $\tau_{\max}$, α can be set to $1/\tau_{\max}$, where $\tau_{\max}$ is typically in the order of seconds or minutes (e.g., 50 seconds). Note that in one embodiment, in the optimization framework, users do not need to wait until a full scanning is completed and/or necessary patches are applied. If the risk they impose on the network as captured by Eq. 1 is acceptable and there is greater gain in admitting the user earlier rather than after performing a full scan, then the user can be admitted earlier. However, the scanning process and the reputation updates continue even after the user is admitted to the network.

[0046] In one embodiment, the admission control policy defines the damage $D(B_i)$ of subset B_i as a linear function of sum rate of B_i, i.e., $D(B_i) = s \times \Sigma_i$. In one embodiment, s is set to 1. In another embodiment, instead of using the exact value of the expected value computation in (1), an upper bound that is convex in p_i's can be used. One such function is in the form of

$$E_B[D] = \gamma \sum_{i=1}^{N} (1 - p_i) r_i$$

with γ being a real value between 1 and 2^N, r_i is the allowed connection rate of i-th user. Note that γ^{2N} is an upper bound. Also, smaller values of γ can be used to more tightly upper bound the expected damage.

[0047] With the linear approximations for the constraints, the optimization problem becomes a convex optimization problem. Accordingly, the following Lagrangian function is defined:

$$\phi(\lambda, \mu) = \sum_{i=1}^{N} U_i(\tau_i) - \lambda \left(\gamma \sum_{i=1}^{N} (1 - p_{i0} - \alpha \tau_i) \tau_i - \Gamma_{th} \right) - \sum_{i=1}^{N} \mu_i (p_{i0} + \alpha \tau_i - 1)$$

where λ and μ_i are the Lagrange multipliers. Solving for the Kuhn-Tucker conditions reveals that:

$$\sum_{i \in A} U_i'^{-1}[-\alpha \gamma r_i \lambda] r_i = \frac{1}{\alpha} \left(\sum_{i \in A} (1 - p_{i0}) r_i - \frac{\Gamma_{th}}{\gamma} \right) \quad (9)$$

where $U_i'(.)$ is the derivative of U_i and $U_i'^{-1}(.)$ is the inverse function of $U_i'(.)$.

[0048] Let A be the set of users who are admitted into the system before they reach the maximum reputation level. Once λ is computed, the admission delay of each user i can be computed as:

$$\tau_i = \begin{cases} \tau_{i,\max}; i \notin A \\ U_i'^{-1}[-\alpha\gamma r_i \lambda]; i \in A \end{cases} \quad (10)$$

Here, $\tau_{i,\max} = (1 - p_{i0})/\alpha$ represents the maximum delay user i can observe before being admitted into the system (after $\tau_{i,\max}$ the user's reputation becomes one, thus it does not pose a security risk).

[0049] To find the admission times for each user, the value of the set A needs to be determined. First, observe that for each $i \in A$, the following inequality exists:

$$\tau_i = U_i'^{-1}[-\alpha\gamma r_i \lambda] < \tau_{i,\max}$$

or equivalently,

$$\lambda < \frac{-U_i'(\tau_{i,\max})}{\alpha\gamma r_i} = \lambda_i \quad (11)$$

due to the fact that $U_i'(x)$ is a decreasing function of x. For purposes herein, λ_i is referred to as the user priority function. Note that the lower λ values would result in admitting more users into the system since more users would satisfy the above inequality. A lower λ would also decrease the admission delay of each user since $U_i'(x)$ is a decreasing function of x and so $U_i'^{-1}(-x)$ is an increasing function of x. Therefore, a lower λ value implies a lower admission delay for users in set A. As a result, the sum utility increases. In one embodiment, the optimum solution finds the smallest $\lambda(A) > 0$ such that the above constraints are satisfied.

[0050] Note that if a user with λ_i is in A, then all j such that $\lambda_j \geq \lambda_i$ must be in A. In one embodiment, all users are sorted in increasing λ_i such that $\lambda_{j1} \leq \lambda_{j2} \leq \ldots \leq \lambda_{jN}$. Set A can then be one of the following $(N+1)$ subsets $A(0) = \{j_1, \ldots, j_N\}$, $A(1) = \{j_2, \ldots, j_N\}$, $A(2) = \{j_3, \ldots, j_N\}$, …, $A(N-1) = \{j_N\}$, $A(N) = \emptyset$. In one embodiment, the admission controller executes the following algorithm:

for m=0 to N

$A := A(m)$;

Compute λ *according to Eq. (9)*;

Set $A' = \{\forall i : \lambda_i > \lambda\}$;

If $A' \equiv A$ *set* $\lambda^* = \lambda$ *and exit;*

Else continue to next iteration;

end

[0051] After the algorithm halts, the admission controller computes the admission time τ_i of each user *i* according to Eq. (10) using $\lambda = \lambda^*$ and the last *A*. The users are then admitted at time $\tau_{arr,i} + \tau_i$, where $\tau_{arr,i}$ is the arrival time of user *i* into the system.

[0052] In a typical case, where there are multiple arrivals of users over a certain period of time, each new user arrival triggers the re-execution of the algorithm for the users who are waiting for admission either because their admission time is not reached or because their admission time is not yet computed (e.g., the newly arrived users). Suppose *Z* denotes the set of users at time *t* who are already admitted into the network, not yet departed, and have reputation less than one, i.e., $Z = \{\forall i : \tau_{arr,i} + \tau_i < t < \tau_{depart,i} \wedge p_i < 1\}$. Also, let Ψ be the set of all possible subsets of *Z*. The first constraint given by (5) can then be rewritten as $E_B[D] \le \Gamma_{th} - E_\Psi[D]$. In other words, the damage threshold is reduced by the expected damage that can be caused by different subsets of users who are already admitted into the network. As before when linear approximation is used, in one embodiment, this damage is computed as

$$E_\Psi[D] = \gamma \sum_{i \in Z} (1 - p_i) r_i \quad (12)$$

After the new conditions are taken into account (i.e., the expected damage is updated with the newly arrived users, their reputations, and traffic injection rates), the admission times for the users who are still waiting to enter the network are computed/recomputed using the same algorithm as before.

[0053] Figure 5 is a flow diagram of one of embodiment of the admission control process as described above.

Solution for a Specific Utility Function

[0054] In one embodiment, the utility function $U_i(\tau_i)$ is a concave non-increasing function of τ_i. In one embodiment, the following utility function is used:

$$U_i(\tau_i) = \beta\left(1 + \ln\left(1 - \frac{\tau_i}{\tau_{\max}}\right)\right) \quad (13)$$

where τ_i represents the additional delay due to the re-executed admission control algorithm.

[0055] Using this expression for the utility, the exact values of priority λ_i are calculated from Eq. (11):

$$\lambda_i = \frac{\beta}{\alpha\gamma r_i\left(\tau_{\max} - \tau_{i,\max}\right)} \quad (14)$$

where $\tau_{\max}$ represents the maximum tolerable access delay. In one embodiment, the value of β is 1.

[0056] For this specific utility function, the following observations are evident. First, the users with a higher upload rate have lower λ_i values, and hence they are more likely to be excluded from set A and wait until the end of a full scan. Also, users with lower $\tau_{i,\max}$ (i.e., with higher initial reputation) are also likely to be delayed until their systems are fully scanned. At first sight, this might seem surprising, since the users with a higher initial reputation value are likely to pose less security risk. However, the key here is the utility function: when users have high initial reputation value, their maximum delays $\tau_{i,\max}$ are already small and their utilities are little impacted by the extra delay. Hence, from the network point of view, it is reasonable to preferentially opt for eliminating the security risks of lower reputation users with full scanning and patching.

[0057] Following the steps of the admission algorithm outlined above and utilizing the expression for λ_i from Eq. 14, the expression for λ is computed from Eq. 9:

$$\lambda(A(m)) = \frac{\beta}{\frac{1}{|A(m)|}\sum_{i\in A(m)}\frac{\beta}{\lambda_i} + \frac{\Gamma_{th}}{|A(m)|}} \quad (15)$$

and the corresponding λ^*, as it was outlined above, is computed.

[0058] Since the utility function is of form Eq. 13, we have:

$$U_i'^{-1}\left[-\alpha\gamma r_i\lambda^*\right] = \tau_{\max} - \frac{\beta}{\alpha\gamma r_i\lambda^*} \quad (16)$$

and the access delay is:

$$\tau_i = \begin{cases} \tau_{i,\max}; i \notin A \\ \tau_{\max} - \dfrac{\beta}{\alpha \gamma r_i \lambda^*}; i \in A \end{cases} \quad (17)$$

An Example of a Computer System

[0059] Figure 4 is a block diagram of an exemplary computer system that may perform one or more of the operations described herein. Referring to Figure 4, computer system 400 may comprise an exemplary client or server computer system. Computer system 400 comprises a communication mechanism or bus 411 for communicating information, and a processor 412 coupled with bus 411 for processing information. Processor 412 includes a microprocessor, but is not limited to a microprocessor, such as, for example, Pentium™, PowerPC™, Alpha™, etc.

[0060] System 400 further comprises a random access memory (RAM), or other dynamic storage device 404 (referred to as main memory) coupled to bus 411 for storing information and instructions to be executed by processor 412. Main memory 404 also may be used for storing temporary variables or other intermediate information during execution of instructions by processor 412.

[0061] Computer system 400 also comprises a read only memory (ROM) and/or other static storage device 406 coupled to bus 411 for storing static information and instructions for processor 412, and a data storage device 407, such as a magnetic disk or optical disk and its corresponding disk drive. Data storage device 407 is coupled to bus 411 for storing information and instructions.

[0062] Computer system 400 may further be coupled to a display device 421, such as a cathode ray tube (CRT) or liquid crystal display (LCD), coupled to bus 411 for displaying information to a computer user. An alphanumeric input device 422, including alphanumeric and other keys, may also be coupled to bus 411 for communicating information and command selections to processor 412. An additional user input device is cursor control 423, such as a mouse, trackball, trackpad, stylus, or cursor direction keys, coupled to bus 411 for communicating direction information and command selections to processor 412, and for controlling cursor movement on display 421.

[0063] Another device that may be coupled to bus 411 is hard copy device 424, which may be used for marking information on a medium such as paper, film,

or similar types of media. Another device that may be coupled to bus 411 is a wired/wireless communication capability 425 to communicate to a network (via a network interface) or another device (e.g., mobile device).

[0064] Note that any or all of the components of system 400 and associated hardware may be used in the present invention. However, it can be appreciated that other configurations of the computer system may include some or all of the devices.

[0065] 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 for admitting a user device waiting to enter a network having one or more currently admitted user devices, the method comprising:

receiving a request from the user device to access the network;

determining whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, wherein the constraint optimization is based on a utility associated with admitting the user device to the network, a reputation value associated with the user device, and a botnet damage estimation on the network associated with the user device; and

admitting the user device to the network based on results of determining whether to admit the user device.

2. An admission control for use in admitting a user device waiting to enter a network having one or more currently admitted user devices, the admission control comprising:

a memory to store instructions; and

a processor coupled to the memory to execute the instructions to perform a method comprising

receiving a request from the user device to access the network,

determining whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, wherein the constraint optimization is based on a utility associated with

admitting the user device to the network, a reputation value associated with the user device, and a botnet damage estimation on the network associated with the user device, and

admitting the user device to the network based on results of determining whether to admit the user device.

3. An article of manufacture having one or more computer readable storage media storing instructions which when executed by an admission controller of a network cause the admission controller to perform a method comprising:

receiving a request from the user device to access the network;

determining whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, wherein the constraint optimization is based on a utility associated with admitting the user device to the network, a reputation value associated with the user device, and a botnet damage estimation on the network associated with the user device; and

admitting the user device to the network based on results of determining whether to admit the user device.

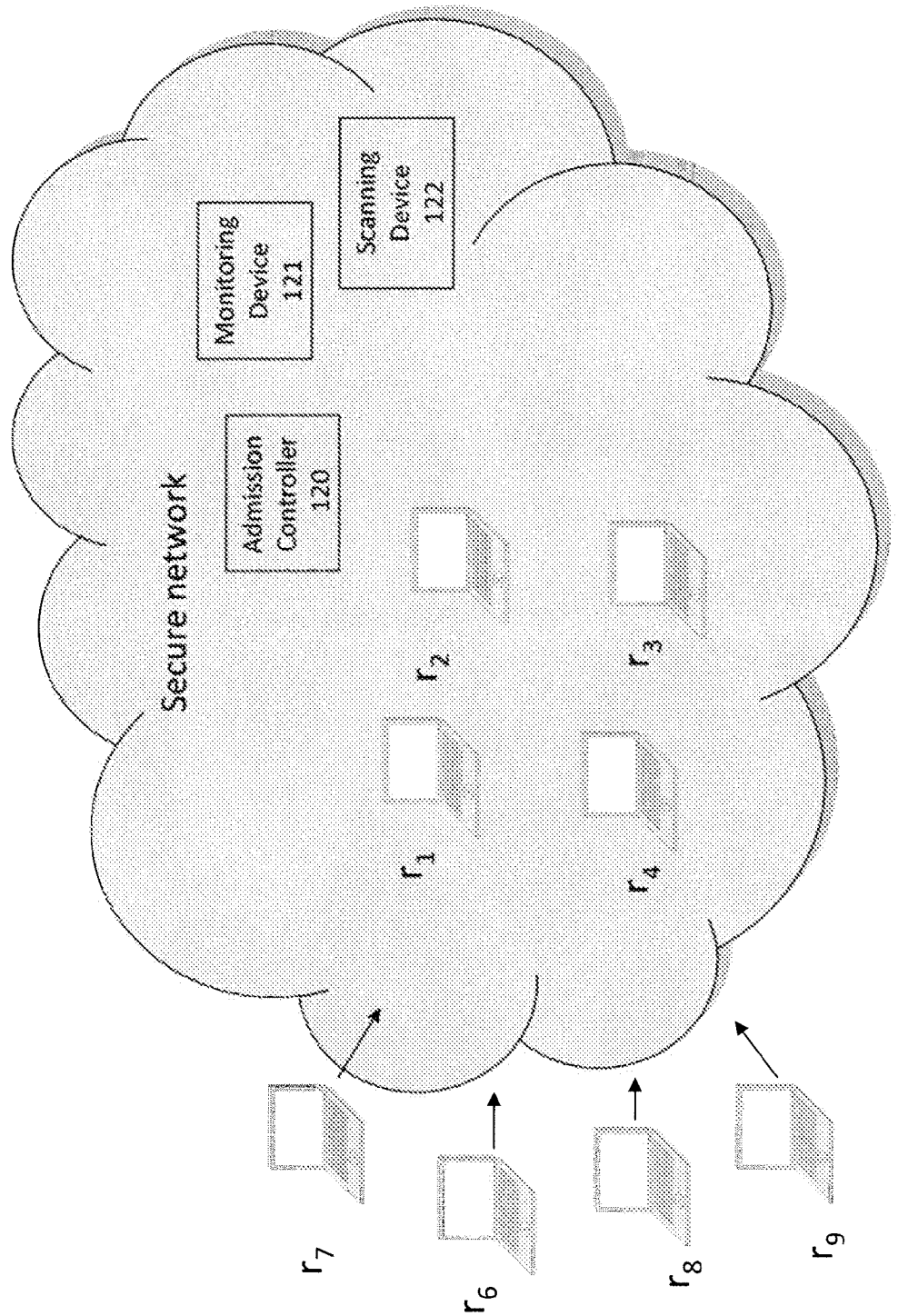
Secure network
Admission Controller 120
Monitoring Device 121
Scanning Device 122
r1
r2
r3
r4
r6
r7
r8
r9

Figure 1

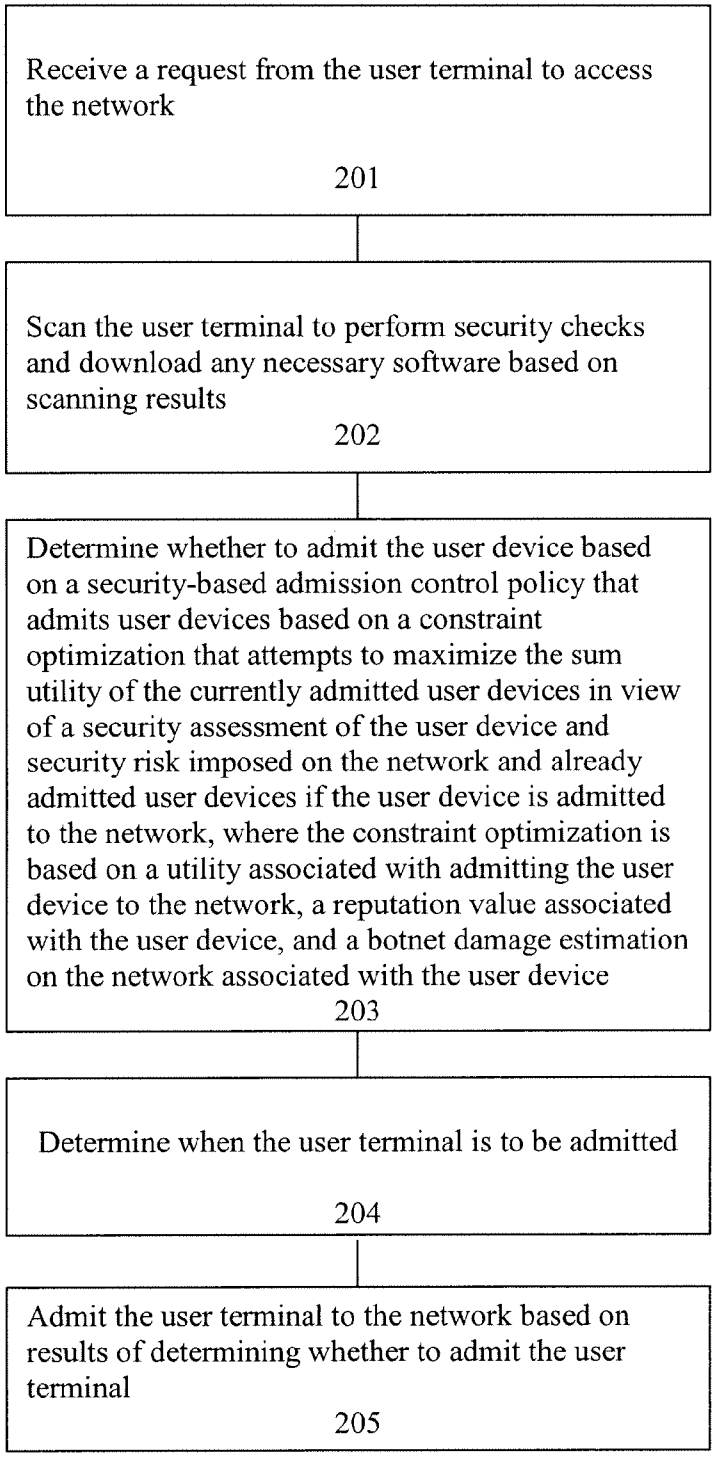
Receive a request from the user terminal to access the network
201
Scan the user terminal to perform security checks and download any necessary software based on scanning results
202
Determine whether to admit the user device based on a security-based admission control policy that admits user devices based on a constraint optimization that attempts to maximize the sum utility of the currently admitted user devices in view of a security assessment of the user device and security risk imposed on the network and already admitted user devices if the user device is admitted to the network, where the constraint optimization is based on a utility associated with admitting the user device to the network, a reputation value associated with the user device, and a botnet damage estimation on the network associated with the user device
203
Determine when the user terminal is to be admitted
204
Admit the user terminal to the network based on results of determining whether to admit the user terminal
205

Figure 2

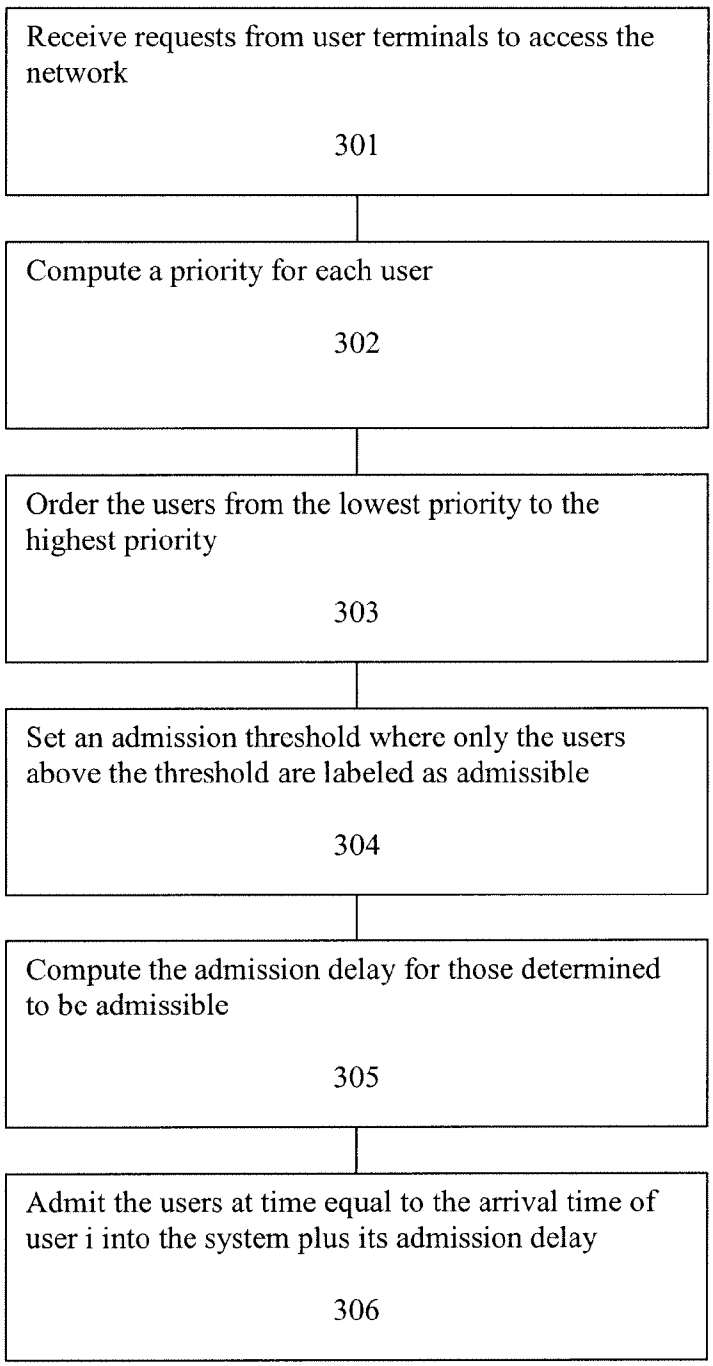
Receive requests from user terminals to access the network
301
Compute a priority for each user
302
Order the users from the lowest priority to the highest priority
303
Set an admission threshold where only the users above the threshold are labeled as admissible
304
Compute the admission delay for those determined to be admissible
305
Admit the users at time equal to the arrival time of user i into the system plus its admission delay
306

Figure 3

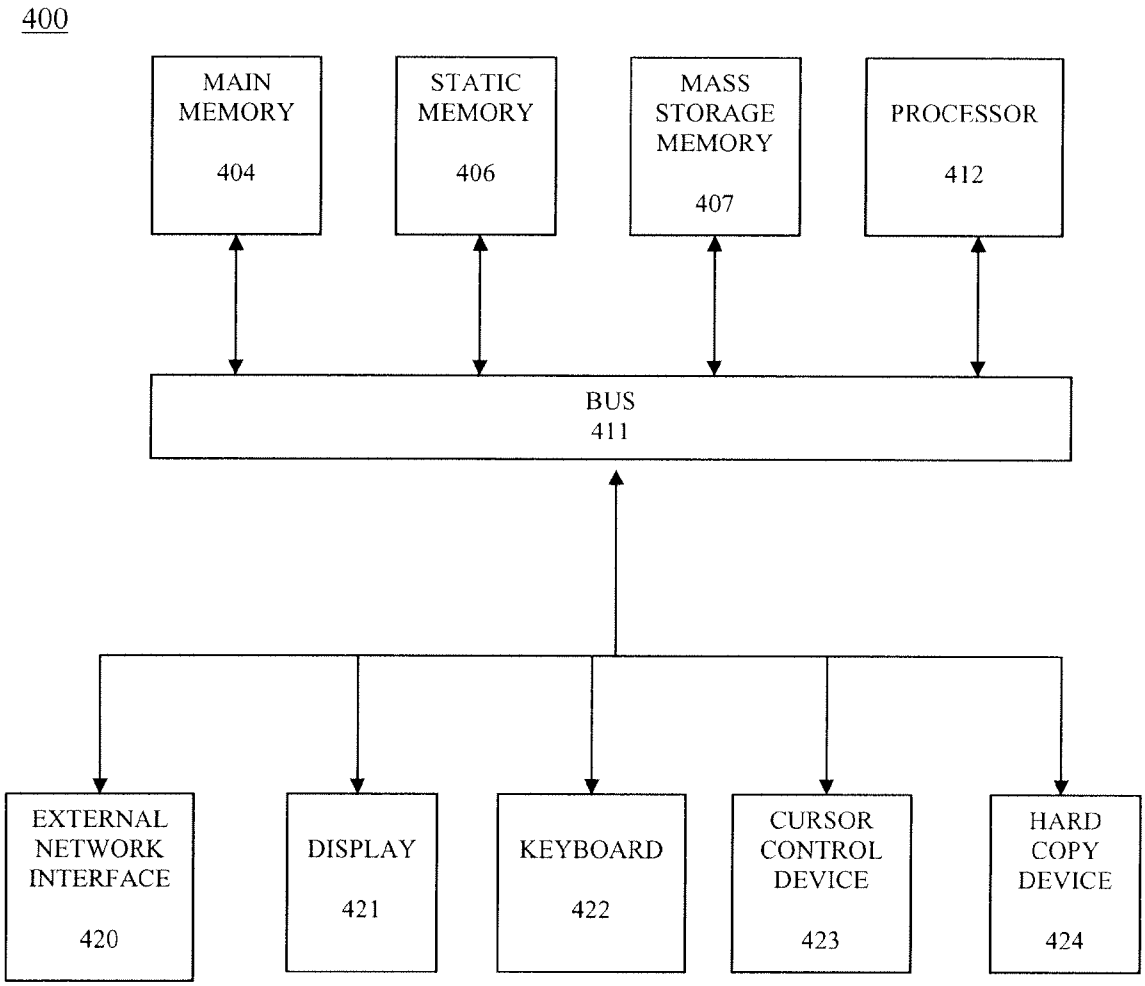
400
MAIN MEMORY 404
STATIC MEMORY 406
MASS STORAGE MEMORY 407
PROCESSOR 412
BUS 411
EXTERNAL NETWORK INTERFACE 420
DISPLAY 421
KEYBOARD 422
CURSOR CONTROL DEVICE 423
HARD COPY DEVICE 424

FIG. 4

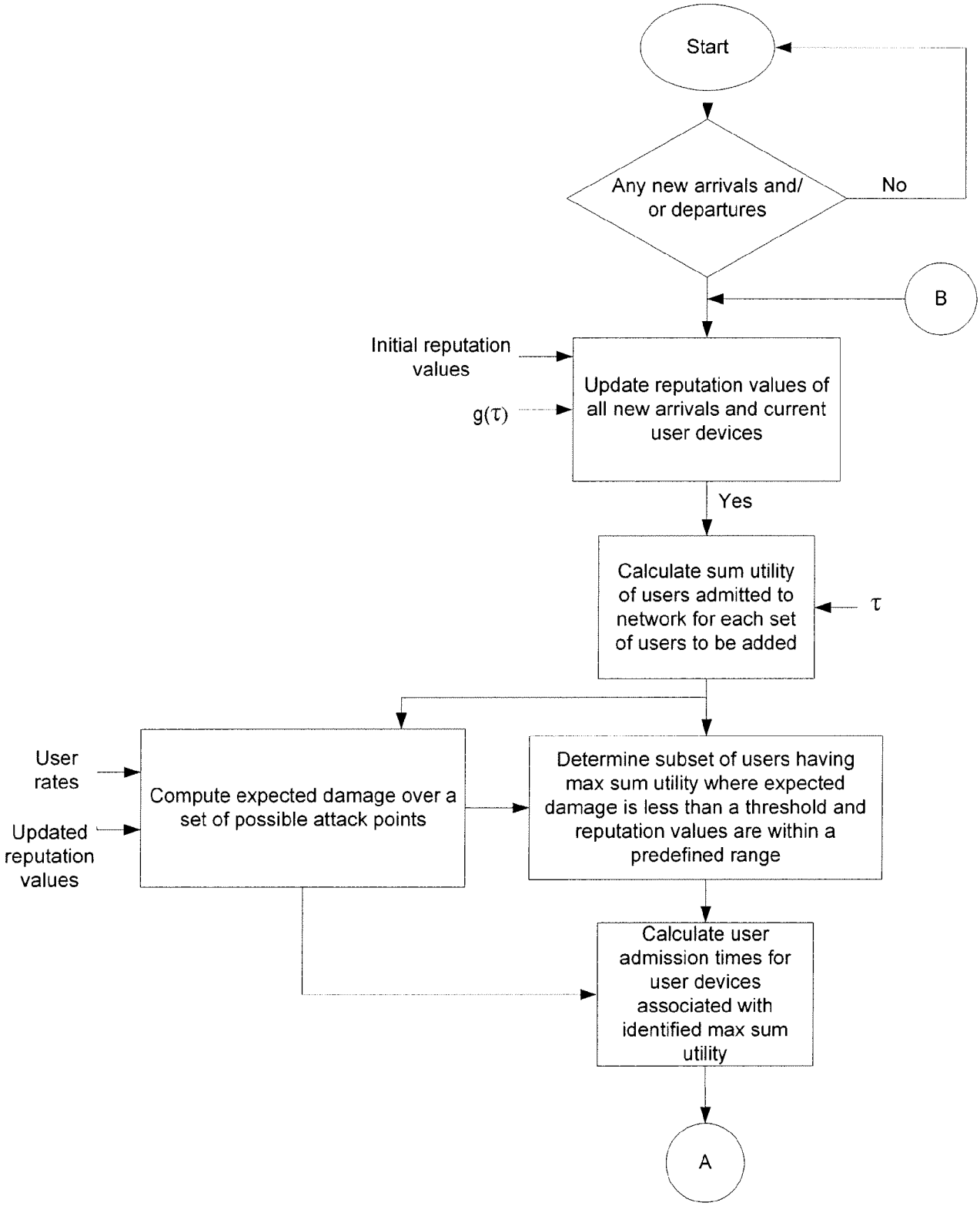
Start
Any new arrivals and/ or departures
No
B
Initial reputation values
g(τ)
Update reputation values of all new arrivals and current user devices
Yes
Calculate sum utility of users admitted to network for each set of users to be added
τ
User rates
Updated reputation values
Compute expected damage over a set of possible attack points
Determine subset of users having max sum utility where expected damage is less than a threshold and reputation values are within a predefined range
Calculate user admission times for user devices associated with identified max sum utility
A

Figure 5

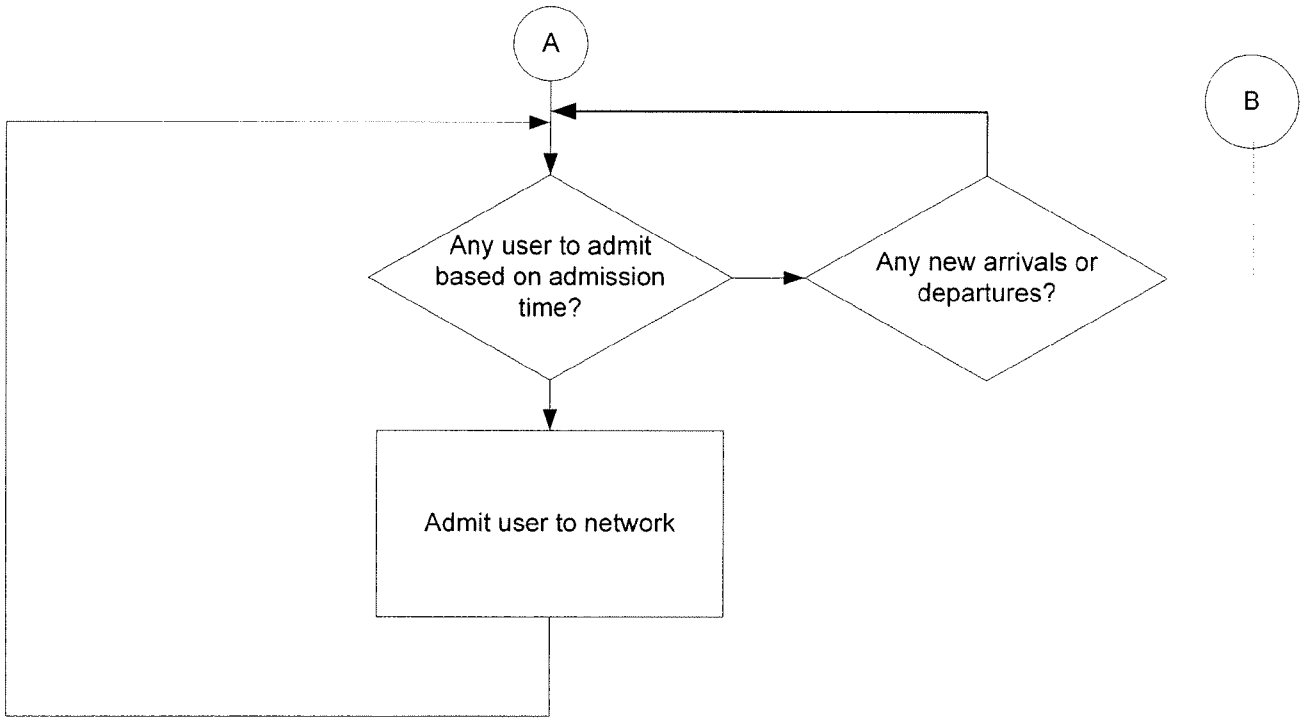

Figure 5 cont'd.

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/057685

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04L29/06 H04L29/08

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

B. FIELDS SEARCHED

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

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/228874 A1 (EDGETT JEFF S [US] ET AL EDGETT JEFF STEVEN [US] ET AL) 13 October 2005 (2005-10-13) abstract figures 4-7 paragraph [0003] - paragraph [0018] paragraph [0024] - paragraph [0027] paragraph [0032] paragraph [0034] - paragraph [0047] paragraph [0059] - paragraph [0088]	1-3
X	WO 2005/111841 A2 (TRUSTED NETWORK TECHNOLOGIES I [US]; SHAY A DAVID [US]) 24 November 2005 (2005-11-24) abstract page 2, line 1 - page 11, line 17 page 16, line 27 - page 18, line 12	1-3

-/--

[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
21 January 2010	28/01/2010

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, Fax: (+31-70) 340-3016	Kopp, Klaus

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2009/057685

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2005/107204 A1 (TELECOM ITALIA SPA [IT]; DE LUTIIS PAOLO [IT]; DI CAPRIO GAETANO [IT];) 10 November 2005 (2005-11-10) abstract page 8, line 10 - page 11, line 15	1-3

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2009/057685

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005228874 A1	13-10-2005	EP 1743252 A2	17-01-2007
		US 2009150525 A1	11-06-2009
		WO 2005104425 A2	03-11-2005
WO 2005111841 A2	24-11-2005	WO 2005114898 A2	01-12-2005
WO 2005107204 A1	10-11-2005	EP 1743465 A1	17-01-2007
		US 2007233883 A1	04-10-2007