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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 October 2010 (28.10.2010)

PCT

(10) International Publication Number
WO 2010/123623 A2

(51) International Patent Classification:
G06F 21/20 (2006.01) *G06F 17/20* (2006.01)
G06F 15/16 (2006.01) *G06F 9/06* (2006.01)

(21) International Application Number:
PCT/US2010/025702

(22) International Filing Date:
27 February 2010 (27.02.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/171,264 21 April 2009 (21.04.2009) US
61/241,389 10 September 2009 (10.09.2009) US
12/709,504 21 February 2010 (21.02.2010) US

(71) Applicant *(for all designated States except US)*: **BRIGHTCLOUD INCORPORATED** [US/US]; 4810 Eastgate Mall, San Diego, CA 92121 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **HEGLI, Ronald** [US/US]; 12855 Chaparral Ridge Road, San Diego, CA 92130 (US). **LONAS, Hal** [US/US]; 908 Orchid Way, Carlsbad, CA 92011 (US). **HARRIS, Christopher** [US/US]; 7465 Healis Place, San Diego, CA 92129 (US).

(74) Agent: **CATANIA, Michael**; Clause Eight IPS, P.O. Box 131270, Carisbad, CA 92013 (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, TH, 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).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: SYSTEM AND METHOD FOR DEVELOPING A RISK PROFILE FOR AN INTERNET RESOURCE

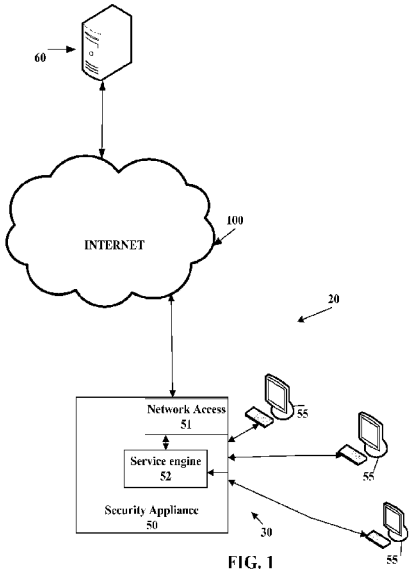


FIG. 1

(57) **Abstract**: A method (1000) and system (20) for controlling access to an Internet resource is disclosed herein. When a request for an Internet resource, such as a Web site, is transmitted by an end-user of a LAN (30), a security appliance (50) for the LAN (30) analyzes a reputation index for the Internet resource at a server (60) before transmitting the request over the Internet (100). The reputation index is based on a reputation vector which includes a plurality of factors for the Internet resource such as country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, to-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency.

WO 2010/123623 A2

System And Method For Developing A Risk Profile For An Internet Resource (Docket Number BC-001PCT)

Technical Field

[0001] The present invention is related to assessing risk profiles of Internet resources. More specifically, the present invention is related to a system and method for developing a risk profile for an Internet resource by generating a reputation index, based on attributes of the resource collectively referred to as the reputation vector of the resource.

Background Art

[0002] Management of internet access, particularly to Web sites, has been accomplished in the past using "Content Filtering", where Web sites are organized into categories and requests for Web content are matched against per-category policies and either allowed or blocked. This type of management focuses on the subject matter of a Web site, and provides visibility into, for example, how employees spend their time, and their company's network bandwidth usage, during the course of the day. These solutions also allow companies to enforce established internet usage policy (IUP) by blocking Web sites whose subject matter violates their IUP.

[0003] Security solutions, such as anti-virus products, examine file or Web page content to discover known patterns or signatures that represent security threats to users, computers, or corporate networks. These focus not on the subject matter of a site, but look for viruses and other 'malware' that are currently infecting the site. However, current solutions to management of Internet resources fail to measure the security risk associated with accessing an Internet resource in a more predictive way, before infections are isolated and signatures are identified and distributed.

[0004] A possible analogy to the reputation of an Internet resource is the credit score of an individual. A Web user would want to be informed of the reputation of a Web site before visiting it, just as a lender would want to know the reputation, the financial reputation at least, of a borrower of the lender's money.

[0005] A credit score is based on a variety of fairly tightly related factors, such as existing debt, available credit lines, on-time payments, existing credit balances, etc.

[0006] In the United States, a credit score is a number based on a statistical analysis of a person's credit files that represents the creditworthiness of that

person, which is the likelihood that the person will pay their bills. A credit score is primarily based on credit information, typically from one of the three major credit agencies.

[0007] There are different methods of calculating credit scores. The best known one, FICO, is a credit score developed by the Fair Isaac Corporation. FICO is used by many mortgage lenders that use a risk-based system to determine the possibility that the borrower may default on financial obligations to the mortgage lender.

[0008] FICO® scores are provided to lenders by the three major credit reporting agencies: Equifax, Experian and TransUnion. When lenders order your credit report, they can also buy a FICO® score that is based on the information in the report. That FICO® score is calculated by a mathematical equation that evaluates many types of information from the borrower's credit report at that agency. In order for a FICO® score to be calculated on the borrower's credit report, the report must contain sufficient information—and sufficient recent information—on which to base a score. Generally, that means the borrower must have at least one account that has been open for six months or longer, and at least one account that has been reported to the credit reporting agency within the last six months.

[0009] FICO scores provide a reliable guide to future risk based solely on credit report data. FICO® scores have a 300–850® score range. The higher the score, the lower the risk. But no score says whether a specific individual will be a "good" or "bad" customer. And while many lenders use FICO® scores to help them make lending decisions, each lender has its own strategy to determine if a potential borrower is a good customer. Although FICO won't reveal exactly how it determines a credit score, it considers the following factors: payment history (35%); outstanding debt (30%); length of credit history (15%); types of credit (10%); and new credit (10%).

[00010] Returning to Internet resources, attackers have been using the Internet to attack the computers and other devices of users of the Internet. Attackers continue to take advantage of flaws in traditional security measures and bypass reputation-based systems to increase attack effectiveness.

[00011] In 2008, massive attacks were conducted that compromised hundreds of thousands of legitimate Web sites with good reputations worldwide with data-stealing malicious code. The attacks included sites from MSNBC, ZDNet, Wired, the United Nations, a large UK government site, and more. In the attacks, when a user's browser opened one of the thousands of compromised sites, a carefully crafted iframe HTML tag redirected users to a malicious site rife with exploits. As a result, malicious code, designed to steal

confidential information, was launched on vulnerable machines. In addition to Web exploits, email spammers are also taking advantage of the reputation of popular email services like Yahoo! and Gmail to bypass anti-spam systems.

[00012] Also, spammers use sophisticated tools and bots to break the "CAPTCHA -" systems that were developed to keep email and other services safe from spammers and other malicious activity. MICROSOFT Live Mail, GOOGLE's popular Gmail service and Yahoo! mail services were all compromised by this breakthrough method. Subsequently, spammers have been able to sign up for the free email accounts on a mass basis and send out spam from email accounts with good reputations. With a free signup process, access to a wide portfolio of services and domains that are unlikely to be blacklisted given their reputation, spammers have been able to launch attacks on millions of users worldwide while maintaining anonymity.

[00013] Thus, prior art solutions have focused on security when accessing known infected sites in the Internet from a network such as a local area network or a wide area network.

[00014] Hegli *et al.*, U.S. Patent Number 7483982 for *Filtering Techniques For Managing Access To Internet Sites Or Other Software Applications* discloses a system and method for controlling an end user's access to the Internet by blocking certain categorized sites or limiting access based on bandwidth usage.

[00015] Hegli *et al.*, U.S. Patent Number 6606659 for a *System And Method For Controlling Access To Internet Sites* discloses a system and method for controlling an end user's access to the Internet by blocking certain categorized sites or limiting the number of times the end user can access an Internet site.

[00016] Yavatkar *et al.*, U.S. Patent Number 6973488 for *Providing Policy Information To A Remote Device* discloses a method for distributing high level policy information to remote network devices using a low-level configuration.

[00017] Turley *et al.*, U.S. Patent Publication Number 2005/0204050 for a *Method And System For Controlling Network Access* discloses a system and method for controlling access to a specific site by using a gateway that assigns incoming traffic to specific sections of the site.

[00018] Shull *et al.*, U.S. Patent Number 7493403 for *Domain Name Validation* discloses accessing domain name registries to determine the ownership of a domain and monitoring the domain and registry.

[00019] Roy *et al.*, U.S. Patent Number 7406466 for a *Reputation Based Search* discloses using a search engine to present search results associated with measures of reputation to overcome the problem of META tags skewing the search results.

[00020] Hailpern *et al.*, U.S. Patent Number 7383299 for a *System And Method For Providing Service For Searching Web Site Addresses* discloses

[00021] Moore *et al.*, U.S. Patent Number 7467206, for a *Reputation System For Web Services* discloses a system and method for selecting a Web service from a search engine list which is ranked based on reputation information for each Web service.

[00022] The prior art fails to provide solutions to the problems with accessing the Internet.

Summary of the Invention

[00023] The present invention provides a predictive approach based on a statistical model built on a broad sampling of Internet resources with varying degrees of risk. The present invention focuses on the reputation of a Web site, or any Internet-based service or resource. The reputation incorporates many factors that are relevant to the overall safety of visiting a site. The reputation assesses the over-time track record of the site and the provider that operates the web site, the current characteristics of the pages and related files composing the site, and reputations of sites linked to the site and of referrers to the site. The overall assessment is expressed as a score, not unlike a FICO score, that predicts the potential risk of visiting the site which can be used to protect users from inadvertently visiting or utilizing higher-risk sites or services within the Internet.

[00024] There are many components of reputation available within the Internet. Much like other scoring mechanisms, such as credit scoring, the factors to be considered must be decided upon, and the weight that each factor will have in the overall "score" must be determined.

[00025] The present invention provides a system and method for defining a reputation of an Internet service such as a Web site.

[00026] A basic element of reputation is how long a domain has been registered to a particular company/entity. In addition, a domain which frequently changes hands is also interesting in a negative way relative to reputation.

[00027] Preferred steps of the invention are: evaluation of the important features to be included in the collection of reputation-relevant features referred to as the reputation vector; collection of the reputation vectors for a large sample of Internet resource; training of a classifier based on training sets of known high and low reputation services/sites; testing of a model against a wide variety of random samples; run-time evaluation of requests for the Internet resource using the developed classifier and responding to reputation

index information requests from clients which enforce network security policy.

[00028] The present invention preferably protects users against threats which are typically are not related to the subject matter of the service, or site. The present invention preferably protects users and networks from zero-day threats which have not been characterized or included in anti-virus signature files. The present invention preferably allows network managers to protect users and infrastructure without having to restrict access to particular categories of content. The present invention preferably allows higher security which is independent of cultural or moral biases related to many categories of content.

[00029] One aspect of the present invention is a method for controlling access to a Web site. The method includes transmitting a request for a Web site from a browser on a client-side device of a local area network. The Web site resides at a first server. The method also includes receiving the request for the Web site at a security appliance of the local area network prior to transmission of the request over the Internet. The method also includes analyzing a reputation vector for the Web site at the security appliance. The reputation vector includes a plurality of factors for the Web site comprising at least one or more of country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency. The method also includes generating a reputation index for the Web site based on the analysis of the plurality of factors. The method also includes determining if the reputation index for the Web site is above a threshold value established for the local area network. The method also includes transmitting a decision transmission to the browser of the client-side device.

[00030] If the reputation index for the Web site is above the threshold value, the method further includes transmitting the request for the Web-site over the Internet to a server for the Web site and receiving a Web page for the Web site at the local area network. In this situation, the decision transmission is the Web page for the Web site. If the reputation index for the Web site is at or below the threshold value, the decision transmission is a Web page from the local area network containing information that accessing the Web site is against an Internet policy established on the local area network.

[00031] The method can further include obtaining the plurality of factors for the Web site. Obtaining the plurality of factors for the Web site comprises accessing the Web site, analyzing a plurality of HTML documents for the Web

site by crawling the Web site. Accessing the Web site comprises rendering a page for the Web site. Analyzing the plurality of HTML documents comprises determining the JavaScript block count and the picture count of each of the HTML documents.

[00032] Another aspect of the present invention is a system for controlling access to a Web site. The system includes a network, a Web site and a local area network. The network is the Internet. The Web site is hosted at a first server and accessible over the Internet. The local area network includes a plurality of client-side devices and a security appliance. Each of the client side devices has a browser. The security appliance controls access to the Internet by each of the plurality of client-side devices. The security appliance has a service engine for analyzing a reputation vector for the Web site and generating a reputation index for the Web site from the reputation vector. The reputation vector is based on a plurality of factors for the Web site. The plurality of factors comprises at least one or more of country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency. Access to the Web site by any of the plurality of client-side devices is determined on the reputation index exceeding a threshold value established for the local area network.

[00033] Another aspect of the present invention is a method for controlling access to an Internet resource utilizing a reputation generating site. The method includes transmitting a request for an Internet resource from a browser for a client-side device of a local area network. The Internet resource resides at a first server. The method also includes receiving the request for the Internet resource at reputation generating site prior to transmission of the request over the Internet to the first server. The method also includes analyzing a reputation vector for the Internet resource at the reputation generating site. The reputation vector includes a plurality of dimensions for the Internet resource comprising at least two of country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, to-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency. The method also includes generating a reputation index for the Internet resource based on the analysis of the plurality of factors. The method also includes determining if the reputation index for the Internet resource is above a threshold value

established for the local area network. The method also includes transmitting a decision transmission to the browser of the client-side device.

[00034] Another aspect of the present invention is a method for controlling access to an Internet resource. The method includes transmitting a request for an Internet resource from an Internet-enabled client application from a client-side device of a local area network. The Internet resource resides at a first server. The method also includes receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet. The method also includes determining if a reputation index for the Internet resource is at or above a threshold value established for the local area network. The reputation index is generated from a reputation vector for the Internet resource. The reputation vector comprises a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location. The reputation index preferably resides in a database file at the security appliance, which is immediately accessible by the security appliance for determining whether or not to allow access to the Internet resource. Alternatively, the reputation index is generated in real-time at a data collection site accessible by the security appliance over the Internet, and the reputation index is forwarded to the security appliance from the data collection site upon request. The method also includes transmitting a decision transmission to the Internet-enabled client application of the client-side device. The decision transmission allows or denies access to the Internet resource.

[00035] Yet another aspect of the present invention is a method for controlling access to an Internet resource. The method includes transmitting a request for an Internet resource from a Web browser for a client-side device of a local area network. The Internet resource resides at a first server. The method also includes receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet. The method also includes constructing a reputation vector for the Internet resource at the security appliance. The reputation vector comprises a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location. The method also includes analyzing the reputation vector to generate a reputation index for the Internet resource based on the analysis of the plurality of factors and the reputation classifier. The method also includes determining if the reputation index for the Internet resource is at or above a threshold value established for the local area network. The method also includes transmitting a decision transmission to the Web

browser of the client-side device. The decision transmission allows or denies access to the Internet resource.

[00036] Yet another aspect of the present invention is a method for building a reputation database for Internet resources. The method includes collecting a plurality of factors for an Internet resource site to populate a reputation vector for the Internet resource to perform reputation analysis of the Internet resource. The method also includes receiving the plurality of factors for the Internet resource at a data collection site. The method also includes constructing a reputation vector for the Internet resource at the data collection site. The reputation vector comprises a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location. The method also includes analyzing the reputation vector to generate a reputation index for the Internet resource based on the analysis of the plurality of factors and the reputation classifier. The method also includes storing the reputation index for the Internet resource at the data collection site. The method also includes transmitting the stored reputation index to a local area network upon request for managing access to the Internet resource.

[00037] The method further includes weighting each of the plurality of factors based on empirical knowledge of each of the plurality of factors. The method further includes obtaining the plurality of factors for the Internet resource using a crawler. Obtaining the plurality of factors for the Internet resource preferably comprises accessing the Internet service, analyzing a plurality of HTML documents for the Internet resource, and crawling a plurality of linked Internet resources of the plurality of HTML documents for Internet resource. Analyzing the plurality of HTML documents preferably comprises determining the JavaScript block count and the picture count of each of the HTML documents, browser hijacking, file downloads and a subject matter.

[00038] Yet another aspect of the present invention is a method for controlling access to an Internet resource. The method includes collecting a first plurality of Internet resource reputation vectors. The method also includes partitioning the first plurality of Internet resource reputation vectors into a plurality of training sets. The method also includes training a maximum entropy discrimination classifier with the plurality of training sets, the maximum entropy discrimination classifier trained for a specific local area network. The method also includes testing the trained maximum entropy discrimination classifier using a second plurality of Internet resource reputation vectors. Each of the second plurality of Internet resource reputation vectors is unknown to the trained maximum entropy discrimination classifier. The method also includes evaluating the tested maximum entropy discrimination

classifier. The method also includes providing feedback to the evaluated maximum entropy discrimination classifier. The method also includes utilizing the reputation index at a local area network for managing access to an Internet resource.

[00039] Preferably, each of the first plurality of Internet resource reputation vectors comprises a plurality of dimensions for the Internet resource comprising security history, legitimacy, behavior, associations and location, and the method further comprises weighting each of the plurality of dimensions.

[00040] Yet another aspect of the present invention is a method for training a MED classifier for controlling access to an Internet resource. The method includes collecting a plurality of reputation vectors for Internet resources. The method also includes partitioning the plurality of reputation vectors into training sets. The method also includes training a MED classifier with the training sets. The method also includes testing the trained MED classifier against unknown Internet resources. The method also includes evaluating the trained MED classifier. The method also includes determining if the trained MED classifier has been adequately trained.

Brief Description of the Drawings

[00041] FIG. 1 is a block diagram of a system for controlling access to a Web site.

[00042] FIG. 2 is a block diagram of a system for controlling access to a Web site.

[00043] FIG. 3 is a flow chart of a method for controlling access to a Web site.

[00044] FIG. 4 is a Web page for a requested Web site.

[00045] FIG. 5 is a page for a local area network informing a requestor of the denial of access to a Web site.

[00046] FIG. 6 is a block diagram of an Internet resource having a HTML document that is accessed by a crawler.

[00047] FIG. 7 is a flow chart of a method for generating a reputation index.

[00048] FIG. 8 is a flow chart of a method for controlling access to an Internet resource.

[00049] FIG. 9 is a flow chart of a method for utilizing a MED classifier for controlling access to an Internet resource.

[00050] FIG. 10 is flow chart of a method for controlling access to an Internet resource utilizing a MED classifier.

[00051] FIG. 11 is a block diagram of a system for utilizing a MED classifier for controlling access to an Internet resource.

Best Mode(s) For Carrying Out The Invention

[00052] Reputation is a qualitative assessment of the safety of a website, expressed as a quantitative value that can be used in managing internet usage. Internet resources, such as Web sites, are safe, or of high reputation, if the Internet resource preferably has: a reputable ownership and registration; a consistent history; had consistent content during that history; associated with other high reputation sites; from a geographically safe region; the Internet service provider ("ISP") is well-known and reputable; not been known to be a source of malware infection; and worked cooperatively with the end-user and the end-user's Web browser application.

[00053] While security threats are transitory since they come up suddenly and are mitigated as quickly as possible, reputation is built up over a period of time and is a more enduring quality. Reputation can be lost, or become 'bad', over a period of time with repeated security events, bad associations, and bad behavior. For that reason, the occurrence of a single security breach (the site gets hacked and is a danger to visitors) does not dramatically lower the reputation of a site. Repeated occurrences over time, however, will destroy the reputation of the site.

[00054] Competitive reputation products include social considerations in their definitions, such that a highly reputable site, a site "held in high regard", preferably has these characteristics: established record of Web presence; not a source of network security risk; no introduction of malware; no popup ads; no persistent ad infection; is not pornographic or obscene; and has no illegal content.

[00055] The reputation of an Internet resource is preferably determined by security, legitimacy, behavior, geographic location, associations and additional factors. Legitimacy is determined by the top-level domain, the investment in the Internet resource (virtual hosting with non-affiliated sites, multiple hosting and SSL security), the traffic volume, the category age and the popularity rank. Legitimacy is also preferably determined by any or all of the following: the consistency between the registering and hosting city region or country; and city, region or country associated with the IP address. Behaviors include the use of popup ads, browser hijacking and the use of auto-redirecting. Associations include the number of sites linking into the site, the reputations of the linked in sites and the reputations of the linked-to sites. The geographic location includes the hosting country, the registration country, the region and the city. The geographic location also preferably includes the

consistency between the registering and hosting country and the country associated with the IP address.

[00056] In a most preferred embodiment discussed below, machine learning technologies are utilized for controlling access to an Internet resource. A variation on support vector machine techniques called Maximum Entropy Discrimination ("MED") is a preferred machine learning technology. MED allows a computer to be trained to recognize the relative reputation of an Internet resource based on the features of the Internet resource. The set of features which characterize the reputation of the Internet resource is its reputation vector. Once trained, the computer uses the reputation vector for a requested Internet service to evaluate its reputation index, a score which can be used with empirically developed threshold values to block access where the reputation index is deemed to be too low to be safe.

[00057] A predictive security assessment for an Internet resource is provided based on known facts about the Internet resource, which is more secure than relying only on knowledge of previously experienced security attacks.

[00058] The system preferably provides classification of each Internet resource at run-time given a Uniform Resource Identifier (URI) and the reputation vector of the Internet resource. The system returns a score, or index, expressing the results on a relative scale for use by requesting clients, typically a security product which integrates the reputation assessment as a service.

[00059] The reputation vector preferably comprises a combination of some or all of the following: country of domain registration; country of service hosting; country of IP Address block; age of domain registration; time known to the assessor site; subject matter; classification age (time since last re-categorization); rank (popularity); IP Address; virtual hosting; number of hosts; top-level domain (.com, .biz, .ru, etc); security history; run-time behaviors; popup ads; downloadable executables; virus-infected executables; JavaScript block count; picture count; immediate redirect; and response latency. These features are collected and evaluated for all model training samples and at run-time on a per-user-request basis. Those skilled in the pertinent art will recognize that other factors may be utilized which are relevant to the security as determined by an assessor.

[00060] As shown in FIG.1, a system for controlling access to an Internet service is generally designated 20. The system 20 preferably comprises a local area network 30, the Internet 100 and an Internet service located at a remote server 60. The Internet resource is preferably a Web site. A local area network 30 preferably comprises a security appliance 50 and a plurality of client-side devices 55. The plurality of client-side devices preferably

comprises desktop computers, laptop computers, personal digital assistants, smartphones and the like. Each of the client-side devices 55 preferably has a Web-browser for accessing the Internet from the client side device 55. The security appliance 50 preferably comprises a network access 51 for permitting access to the Internet from the local area network 30, and a service engine 52 for determining if a requested Internet resource has a reputation index that meets a threshold established for the local area network 30.

[00061] A method 1000 for controlling access to a Web site is shown in FIG. 3. At block 1001, a request for a Web site is transmitted from a browser for a client-side device of a local area network which is received at a security appliance of the local area network prior to transmission of the request over the Internet. At block 1002, a reputation index for the Web site is obtained at the security appliance. The reputation index is calculated from a reputation vector which preferably includes a plurality of factors for the Web site comprising country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, popularity rank, internet protocol address, number of hosts, to-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency. At block 1004, a determination is made if the reputation index for the Web site is above a threshold value established for the local area network. At decision 1005, if the reputation index is not above the threshold, then at block 1006 access to the Web site is denied and a transmission of the denial is sent to the client–side device, preferably as a page 500 as shown in FIG. 5. If at decision 1005 the reputation index for the Web site is above the threshold, then the access to the Web site by the client-side device is permitted by the security appliance, and preferably, as shown in FIG. 4, a Web page 400 is provided to the client-side device.

[00062] An alternative embodiment of the system 20 is illustrated in FIG. 2. The system 20 preferably comprises a local area network 30, the Internet 100, an Internet service located at a remote server 60 and a reputation generating site 70 preferably having a crawler 71 and a database 72. The Internet service is preferably a Web site. A local area network 30 preferably comprises a security appliance 50 and a plurality of client-side devices 55. Each of the client-side devices 55 preferably has a Web-browser for accessing the Internet from the client side device 55. The security appliance 50 preferably comprises a network access 51 for permitting access to the Internet from the local area network 30, and a service engine 52 for determining if a requested Internet service has a reputation index that meets a threshold establish for the local area network 30. The reputation generating site 70 provides reputation indices

to service engine 52 of the security appliance 50. The reputation generating site 70 preferably utilizes the crawler 71 and other means to access Internet resources such as the Internet resource located at Web server 60. The other means preferably includes publicly available data feeds, purchased databases, proprietary database, zone files from WHOIS database.

[00063] A flow chart for a method 2000 for generating a reputation index is shown in FIG. 7. At block 2001, a HTTP request is transmitted from a reputation generating site 70 for an Internet resource. From the HTTP request, a crawler 71 of the reputation generating site accesses the Internet resource. In accessing the Internet resource, as shown in FIG. 6, the crawler 71 preferably accesses at least one HTML document 91 of a plurality of HTML documents of the Internet resource 90. At block 2003, from the HTML documents and links within the HTML documents, the crawler 71 obtains information concerning the Internet resource 90. The reputation vector for the Internet resource 90 is based on some of this information obtained by the crawler 71. At block 2004, the reputation vector for the Internet resource is analyzed at the reputation generating site 70. At block 2005, a reputation index for the Internet resource 90 is generated at the data collection site. At block 2006, the reputation index for the Internet resource 90 is stored in a database 72 of the reputation generating site 70. The reputation for the Internet resource is available to the security appliance as updates or individual requests. At block 2007, the reputation index for the Internet resource 90 is transmitted to a LAN 30 for storage in a service engine 52 of a security appliance 50.

[00064] A flow chart for a method 3000 for controlling access to an Internet resource is shown in FIG. 8. At block 3001, a request for an Internet resource is transmitted from an Internet-enabled client application for a client-side device 55 of a LAN 30. At block 3002, the request is received at a security appliance 50 of the LAN 30 prior to transmission of the request over the Internet 100. At block 3003, a reputation index for the Internet resource is accessed from a database of a service engine 52 of the security appliance 50. The reputation index is based on a reputation vector which includes a plurality of factors for the Internet resource comprising at least two or more of country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, security history, popularity rank, internet protocol address, number of hosts, to-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency. At block 3004, a determination is made if the reputation index for the Internet resource is at or above a threshold value established for the LAN 30. At decision 3005, if the reputation index is

below the threshold value, then at block 3006 access to the Internet resource is denied and a transmission of the denial is sent to the client–side device 55. If at decision 3005 the reputation index for the Web site is at or above the threshold value, then the access to the Internet resource by the client-side device 55 is permitted by the security appliance 50.

[00065] Table One provides a list of the attributes for the reputation vector and a description of each of the attributes.

TABLE ONE

Attribute	Description
Country	2-letter code, 3-letter code or full name of country based on IP block
Top-level Domain	.com, .biz, .org, .gov, etc.
Domain Age	Number of months in existence on zone lists, or no less than the classification age
Database Age	Months since Authority was entered into database
Classification Age	Months that the Authority has held its current classification
Hosts	Number of IP's associated with the Authority
Virtually Hosted	T/F if the other authorities share associated IP's
Popups	T/F if the page opens new browser windows on its own
Hijack	T/F does the default page alter the browser configuration
JavaScript	Count of <SCRIPT> blocks in default pages
Executables	T/F does the authorities download executables to client
Pictures	Count of pictures on default page
Latency	Number of milliseconds to return default page
Rank	Numerical ranking, used as T if <2,000,000, F otherwise in modeling
Infected	T/F were infected download files found by AV tools during site analysis
Security Trend	Number of malware infections in past 12 months
Total Security Count	Total number of malware infections known
Redirect	Authority redirects to another Authority
IP Address	Analysis of IP address for known threat sources, reserved IP ranges, and legacy IP address assignments
ISP	Internet service provider
City	
Region	

[00066] Table Two is an example of a "good" Internet resource.

TABLE TWO

Attribute	Value
Authority	USmoney. gov
Country	USA
Top Level Domain	Gov
Domain Age	18
Hosts	2
Virtual Hosts	0
Rank	1
Infected	0
Security Events	0
Recent Events	0
PublicCoIP	0
GovernmentIP	1
Hijack	0
JavaScript	0
Executables	0
Pictures	0
Latency	0
Redirect	0

[00067] Table Three is an example of a "bad" Internet resource.

TABLE THREE

Attribute	Value
Authority	www. c.reditcan.cn
Country	CN
Top Level Domain	CN
Domain Age	3
Hosts	1
Virtual Hosts	1
Rank	0
Infected	0
Security Events	1
Recent Events	1
PublicCoIP	0
GovernmentIP	0
Hijack	0
JavaScript	13
Executables	1
Pictures	14
Latency	826
Redirect	0

[00068] Depending on the threshold value established by the administrator of the LAN, the Internet resource of www. c.reditcan .cn with an reputation index value of 51, is not available for access by a user based on its reputation index, and the Internet resource of www. USmoney .GOV is available for access by a user based on its reputation index 95. Thus, even if the Internet resource of www. c.reditcan .cn is not a known source of malware or viruses, the present

invention would prevent an end user client from accessing the Internet resource since its reputation index is deemed unsafe.

[00069] Another embodiment uses a MED algorithm to build a statistical model on a Web page based on good and bad Internet samples. This embodiment uses a unique optimization algorithm for training, as well as two other optimization steps for calibrating the outputs to be probabilities, in a process that tolerates some input errors while still yielding reliable outputs. Training process feedback loops guide the implementer to improve the model data through splitting data into sets for holdout, training, and testing guided by two criteria: most violating examples, and least understood examples. The implementer using the criteria iteratively improves the quality of the training set which also reduces classifier errors and is exponentially faster than having the implementer manually verify or check the example assignments to categories in random or haphazard order. The examples are randomly reassigned before every training iteration to improve generalization. Sparse matrix math during the classification process improves processing speeds to enable a modest computer to classify millions of URLs per day. The implementation allows for a multiple of dimensions, each representing a fact about the Internet resource, to be included in the reputation model, while classification speed of any particular Internet resource is independent of the number of total dimensions in its reputation vector.

[00070] This embodiment is preferred since classifying a large percentage of existing Web sites into reputation risk assessments quickly and efficiently requires an automated process because the number of humans required is too large to be practical or economical. Further, defining automated classification rules by hand is very hard and requires writing many thousands of extremely specific as well as vague rules. All of these rules will interact with each other in an exponential number of ways, making human-based rule construction a daunting effort as well. The machine learning approach of this embodiment solves the problem by having humans define "training sets" or examples of each topic for the classifier, which then "trains" by optimizing the weights each factor should have in order to reduce classification error the most.

[00071] In addition to providing a good implementation of the learning algorithm, this embodiment efficiently utilizes the human efforts in identifying examples of good and bad reputations.

[00072] This embodiment preferably applies an effective learning formulation based on the principles and theory of MED with an efficient optimization algorithm based on the principles of Platt's sequential minimization optimization ("SMO"), in conjunction with an overall optimization of tunable

parameters and calibrated confidence scores, to solve the learning problem given proper examples of Web sites of good and bad reputation.

[00073] The process then involves having humans examine a list of "most violating" examples, the examples which were marked as being good reputations but received extremely low confidence scores from the classifier (and vice-versa), as well as "least understood" examples, the examples which receive a confidence score close to the prior probability of the reputation.

[00074] By spending human time examining these two classes of examples, the classifier benefits from having egregiously misclassified examples being put into the proper reputation (good or bad) as well as providing the classifier with the largest amount of new information as quickly as possible. This combination improves the classifier's real-world effectiveness very quickly with minimal human effort. Thus, this embodiment efficiently combines human and automated work to solve the problem of automated reputation classification of Internet resources.

[00075] In one method, an evaluation of multiple factors (such as discussed above) is included in determining a reputation vector for an Internet resource. This process is done for multiple Internet resources. Next, reputation vectors for a large sample of Internet resources are collected at a data collection site. Next, a MED classifier is trained using the collection of reputation vectors based on training sets of known high reputation Internet resources and low reputation Internet resources. Next, a MED-based model for classification is tested against a wide variety of random samples of Internet resources. Next, a security appliance is deployed at a LAN. Next, a run-time evaluation of Internet resource requests is performed in using the developed MED classifier for responding to reputation index information requests from clients based on a LAN security policy. The MED-based model for classification is preferably utilized at run-time to calculate a reputation index. In this manner, this embodiment provides a predictive security assessment based on known facts about an Internet resource, which is more secure than relying only on knowledge of previously experienced security attacks. This embodiment provides a LAN real-time updates, real-time classification of non-cached URLs and a real-time feedback loop.

[00076] A flow chart of a method 4000 for utilizing a MED classifier for controlling access to an Internet resource is shown in FIG. 9. At block 4001, multiple reputation vectors for a large sample of Internet resources are collected preferably at a reputation generating site. The reputation vectors for the Internet resources are previously generated as discussed above. At block 4002, the reputation vectors are partitioned into multiple training sets. The

training sets comprise at least two training sets divided into high reputation Internet resources and low reputation Internet resources. At block 4003, a MED classifier is trained using the training sets of high reputation Internet resources and low reputation Internet resources to create a trained MED classifier. At block 4004, the trained MED classifier is tested against a wide variety of Internet resources which are not grouped into training sets and the reputation index is unknown to the trained MED classifier. At block 4005, the tested MED classifier is evaluated to determine the accuracy of the tested MED classifier and to determine the most violating examples of either a wrongly categorized high reputation Internet resource or low reputation Internet resource, and the least understood Internet resources. At decision block 4006, an evaluation of the testing is performed. If the testing was performed correctly, then at block 4007 the MED classifier is considered trained and ready for operations. If the testing was inadequate feedback is provided to the MED classifier concerning the wrongly categorized high reputation Internet resources or low reputation Internet resources, and the least understood Internet resources. The process is continued at block 4003 again until the MED classifier is properly trained.

[00077] In another embodiment, a reputation index is returned immediately from a stored set of reputation indexes calculated prior to the user's request. As shown in FIG. 10, a method for controlling access to an Internet resource utilizing a MED classifier is generally designated 5000. At block 5001, a request for an Internet resource is transmitted from an Internet-enabled client application for a client-side device 55 of a LAN 30. At block 5002, a reputation vector for the Internet resource is analyzed preferably at a MED classifier or at a security appliance for the LAN. At block 5003, a reputation index for the Internet resource is accessed/generated from a database of a service engine 52 of the security appliance 50. The reputation index is preferably based on a reputation vector which includes a plurality of factors for the Internet resource comprising at least two or more of country of domain registration, country of service hosting, country of an internet protocol address block, age of a domain registration, security history, popularity rank, internet protocol address, number of hosts, to-level domain, a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency. At block 5004, a determination is made if the reputation index for the Internet resource is at or above a threshold value established for the LAN 30. At decision 5005, if the reputation index is below the threshold value, then at block 5006 access to the Internet resource is denied and a transmission of the denial is sent to the client–side device 55. If at decision

5005 the reputation index for the Web site is at or above the threshold value, then the access to the Internet resource by the client-side device 55 is permitted by the security appliance 50. In such an embodiment, a pre-calculated reputation index residing on the LAN or quickly available to the security appliance of the LAN provides for a much faster response (if not immediate response) as to the accessibility of the Internet resource.

[00078] FIG. 11 illustrates a system 20 for controlling access to an Internet resource utilizing a MED classifier site 77. The system 20 preferably comprises a local area network 30, the Internet 100, a MED classifier site 77, and an Internet service located at a remote server 60. The Internet resource is preferably a Web site. A local area network 30 preferably comprises a security appliance 50 and a plurality of client-side devices 55. Each of the client-side devices 55 preferably has a Web-browser for accessing the Internet from the client side device 55. The security appliance 50 preferably comprises a network access 51 for permitting access to the Internet from the local area network 30, based on data from the MED classifier site 77, which determines if a requested Internet resource has a reputation index that meets a threshold establish for the local area network 30.

[00079] Table Four provides an example of some dimensions and the sorted model weights of the MED classifier.

TABLE FOUR

Identification	Dimension	Sorted Model weights
1966272070	Domain age	3.785360
2307717	gov	1.969750
1906396306	paris	0.647784
1477426223	Hijack	-19.887100

Claims

1. A method for controlling access to an Internet resource, the method comprising: transmitting a request for an Internet resource from an Internet-enabled client application from a client-side device of a local area network, the Internet resource residing at a first server; receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet; determining if a reputation index for the Internet resource is at or above a threshold value established for the local area network, the reputation index generated from a reputation vector for the Internet resource, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location; and transmitting a decision transmission to the Internet-enabled client application of the client-side device, the decision transmission allowing or denying access to the Internet resource.

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

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

4. The method according to claim 1 wherein the location factor of the plurality of factors for the reputation vector comprises country of domain registration, country of service hosting and country of an internet protocol address block, wherein the legitimacy factor of the plurality of factors for the reputation vector comprises age of a domain registration, popularity rank, internet protocol address, number of hosts, top-level domain, and wherein the behavior factor of the plurality of factors for the reputation vector comprises a plurality of run-time behaviors, JavaScript block count, picture count, immediate redirect and response latency.

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

6. A system for controlling access to a Web site, the system comprising:
a network, the network comprising the Internet;
an Internet resource hosted at a first server and accessible over the Internet; and
a local area network, the local area network comprising
a plurality of client-side devices, each of the client side devices having a browser,
a security appliance, the security appliance controlling access to the Internet by each of the plurality of client-side devices, the security appliance having a service engine for analyzing a reputation vector for the Web site and generating a reputation index for the Web site from the reputation vector, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location;
wherein access to the Internet resource by any of the plurality of client-side devices is determined on the reputation index exceeding a threshold value established for the local area network.

7. The system according to claim 6 wherein the client-side device is a personal computer or a PDA.

8. The system according to claim 6 further comprising a data collection site accessible over the Internet, the security site having a crawler for collecting a plurality of factors from a plurality of Internet resources.

9. The system according to claim 8 wherein the data collection site has a database containing reputation indices for a plurality of Internet resources.

10. A method for controlling access to an Internet resource, the method comprising: transmitting a request for an Internet resource from a Web browser for a client-side device of a local area network, the Internet resource residing at a first server; receiving the request for the Internet resource at a security appliance of the local area network prior to transmission of the request over the Internet; accessing a reputation index for the Internet resource, the reputation index based on a reputation vector for the Internet resource comprising at least one of security history, legitimacy, behavior, associations and location; determining if the reputation index for the Internet resource is at or above a threshold value established for the local area network; and transmitting a decision transmission to the Web browser of the client-side device, the decision transmission allowing or denying access to the Internet resource.

11. A method for building a reputation database for Internet resources, the method comprising: collecting a plurality of factors for an Internet resource site to populate a reputation vector for the Internet resource to perform reputation analysis of the Internet resource; receiving the plurality of factors for the Internet resource at a data collection site; constructing a reputation vector for the Internet resource at the data collection site, the reputation vector comprising a plurality of factors for the Internet resource comprising security history, legitimacy, behavior, associations and location; analyzing the reputation vector to generate a reputation index for the Internet resource based on the analysis of the plurality of factors and the reputation classifier; storing the reputation index for the Internet resource at the data collection site; and transmitting the stored reputation index to a local area network upon request for managing access to the Internet resource.

12. The method according to claim 11 further comprising weighting each of the plurality of factors based on empirical knowledge of each of the plurality of factors.

13. The method according to claim 11 further comprising obtaining the plurality of factors for the Internet resource using a crawler.

14. The method according to claim 13 wherein obtaining the plurality of factors for the Internet resource comprises accessing the Internet service, analyzing a plurality of HTML documents for the Internet resource, and crawling a plurality of linked Internet resources of the plurality of HTML documents for Internet resource.

15. The method according to claim 11 analyzing the plurality of HTML documents comprises determining the JavaScript block count and the picture count of each of the HTML documents, browser hijacking, file downloads and a subject matter.

16. A method for controlling access to an Internet resource, the method comprising: collecting a first plurality of Internet resource reputation vectors; partitioning the first plurality of Internet resource reputation vectors into a plurality of training sets; training a maximum entropy discrimination classifier with the plurality of training sets, the maximum entropy discrimination classifier trained for a specific local area network; testing the trained maximum entropy discrimination classifier using a second plurality of Internet resource reputation vectors, the second plurality of Internet resource reputation vectors unknown to the trained maximum entropy discrimination classifier; evaluating the tested maximum entropy discrimination classifier; providing feedback to the evaluated maximum entropy discrimination

classifier; and utilizing the reputation index at a local area network for managing access to an Internet resource.

17. The method according to claim 16 wherein each of the first plurality of Internet resource reputation vectors comprises a plurality of dimensions for the Internet resource comprising security history, legitimacy, behavior, associations and location, and the method further comprises weighting each of the plurality of dimensions.

18. A method for training a MED classifier for controlling access to an Internet resource, the method comprising: collecting a plurality of reputation vectors for Internet resources; partitioning the plurality of reputation vectors into training sets; training a MED classifier with the training sets; testing the trained MED classifier against unknown Internet resources; evaluating the trained MED classifier; and determining if the trained MED classifier has been adequately trained.

19. The method according to claim 18 were the training set comprises good reputation Internet resources and bad reputation Internet resources.

20. The method according to claim 18 further comprising locating the trained MED classifier at a security appliance of the local area network.

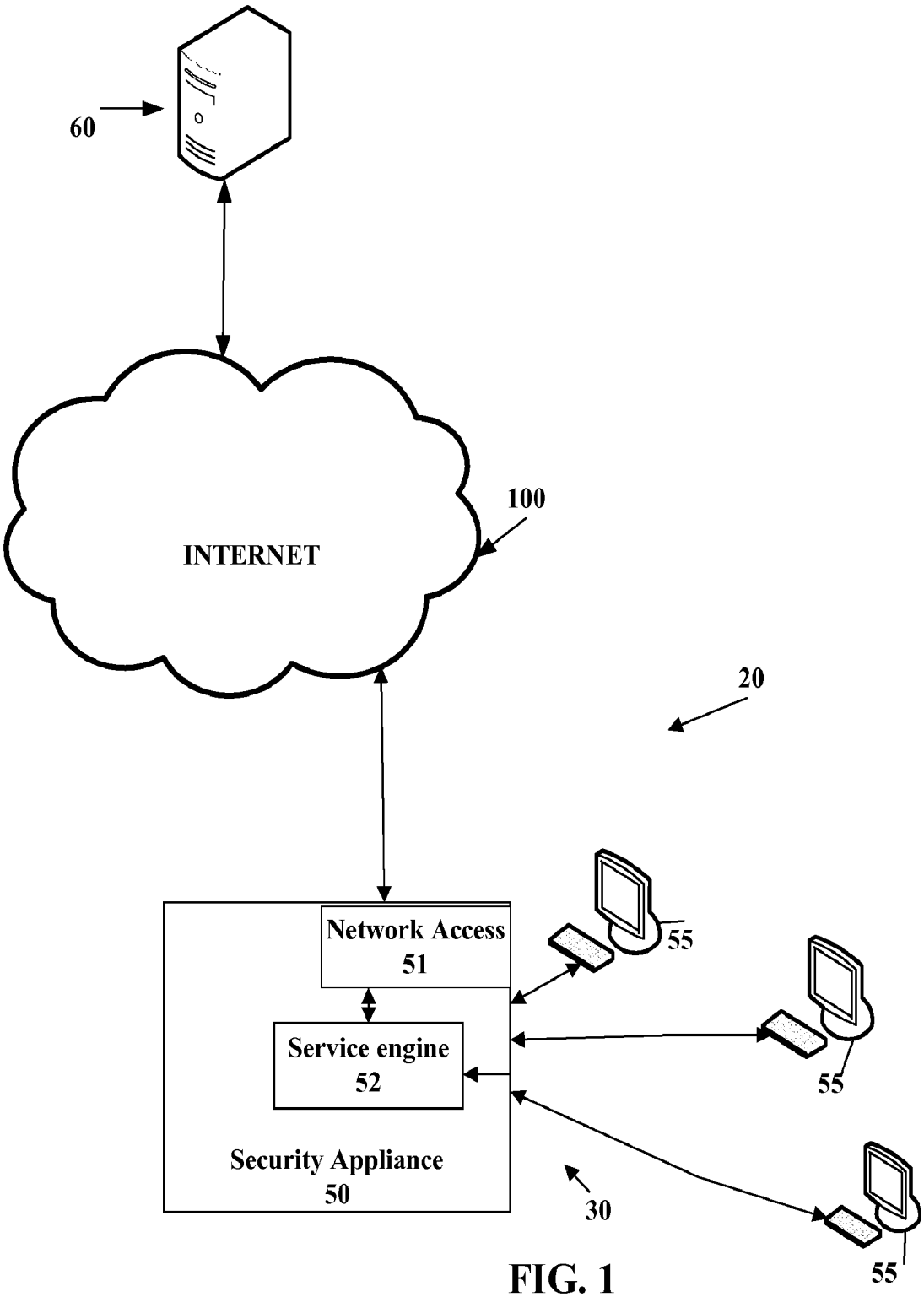

60
INTERNET
100
20
Network Access
51
Service engine
52
Security Appliance
50
30
55
55
55

FIG. 1

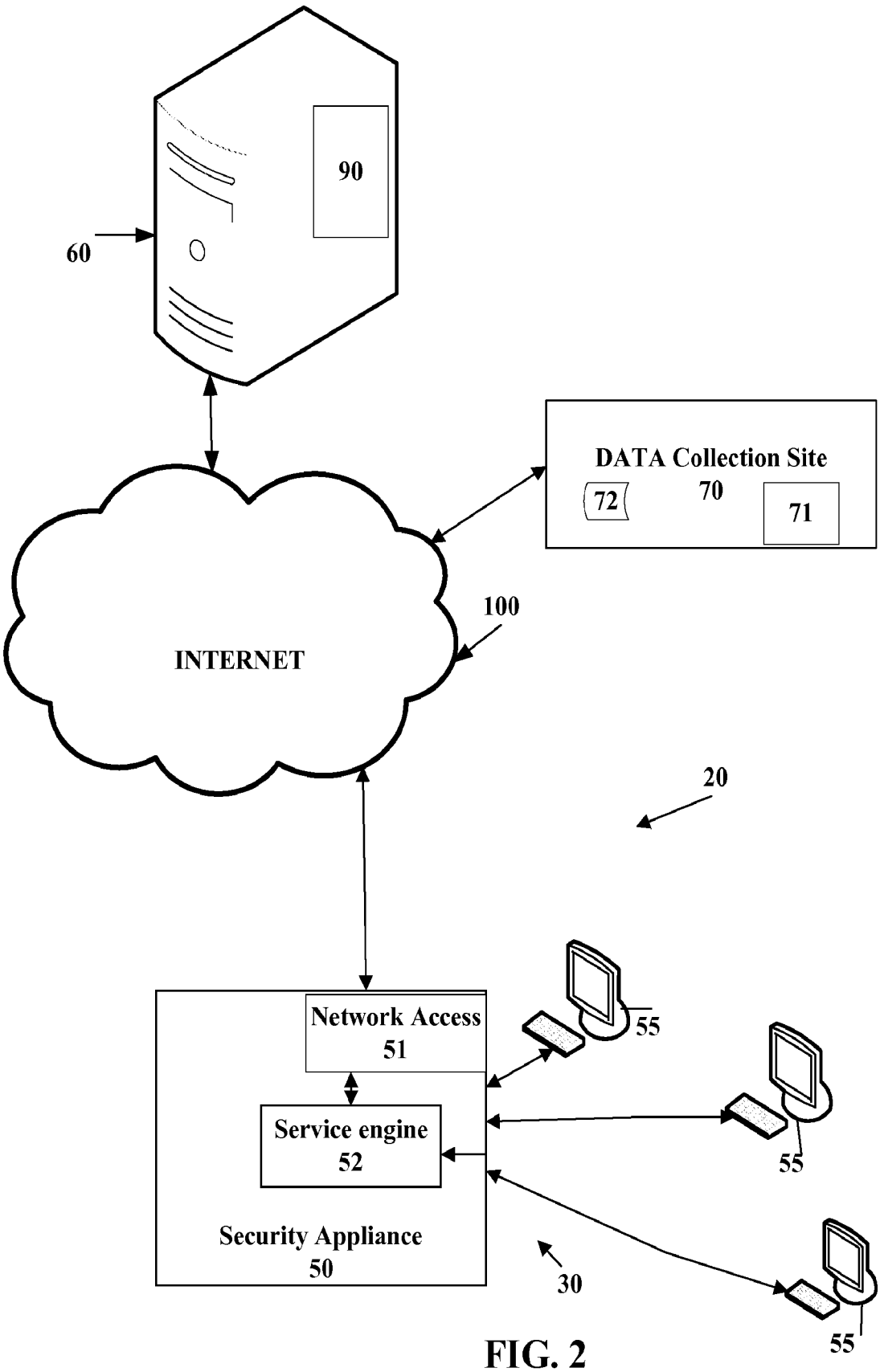

90
60
DATA Collection Site
72
70
71
100
INTERNET
20
Network Access
51
55
Service engine
52
55
Security Appliance
50
30
55

FIG. 2

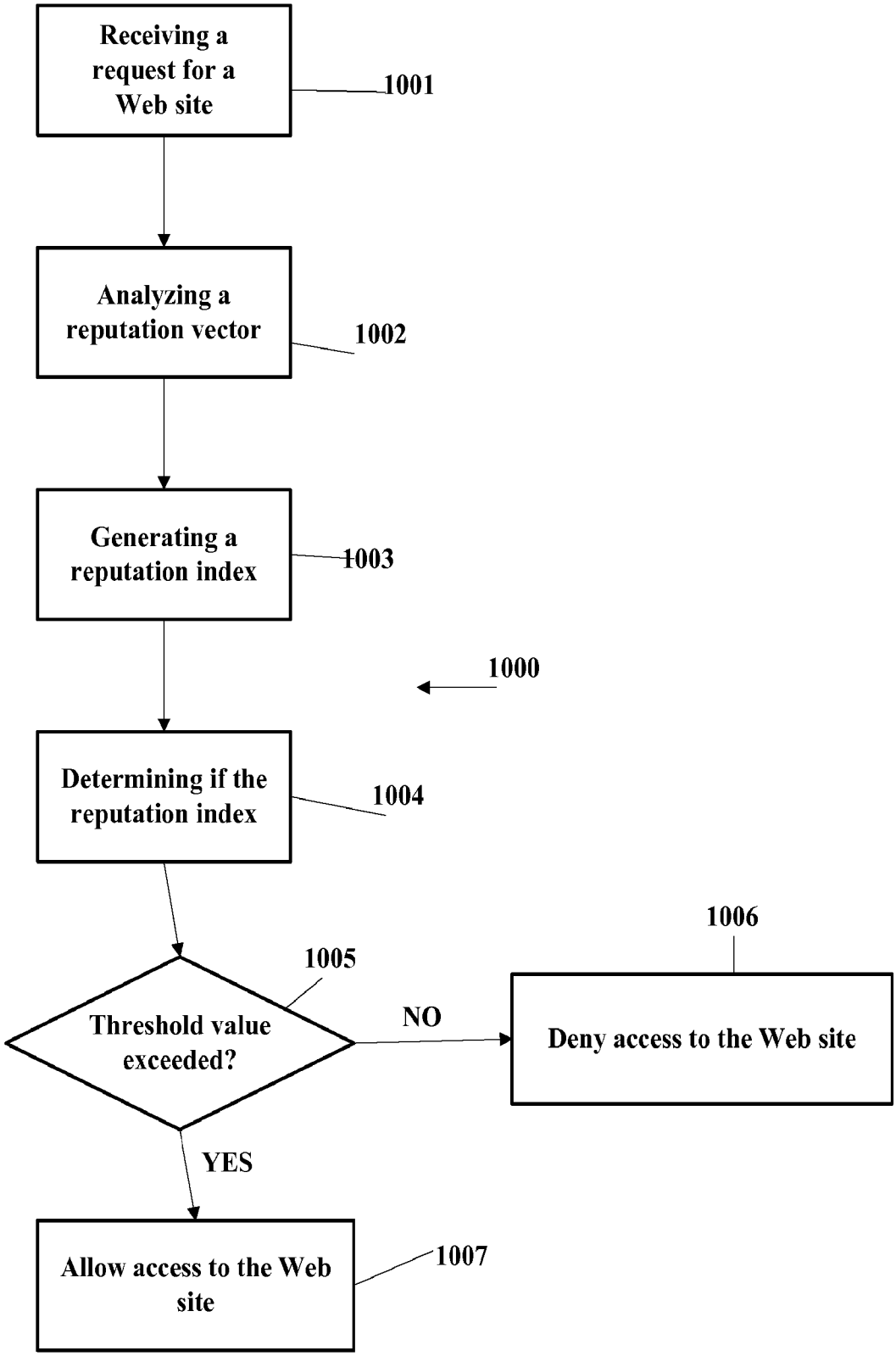

Receiving a request for a Web site
1001
Analyzing a reputation vector
1002
Generating a reputation index
1003
1000
Determining if the reputation index
1004
1005
Threshold value exceeded?
NO
1006
Deny access to the Web site
YES
Allow access to the Web site
1007

FIG. 3

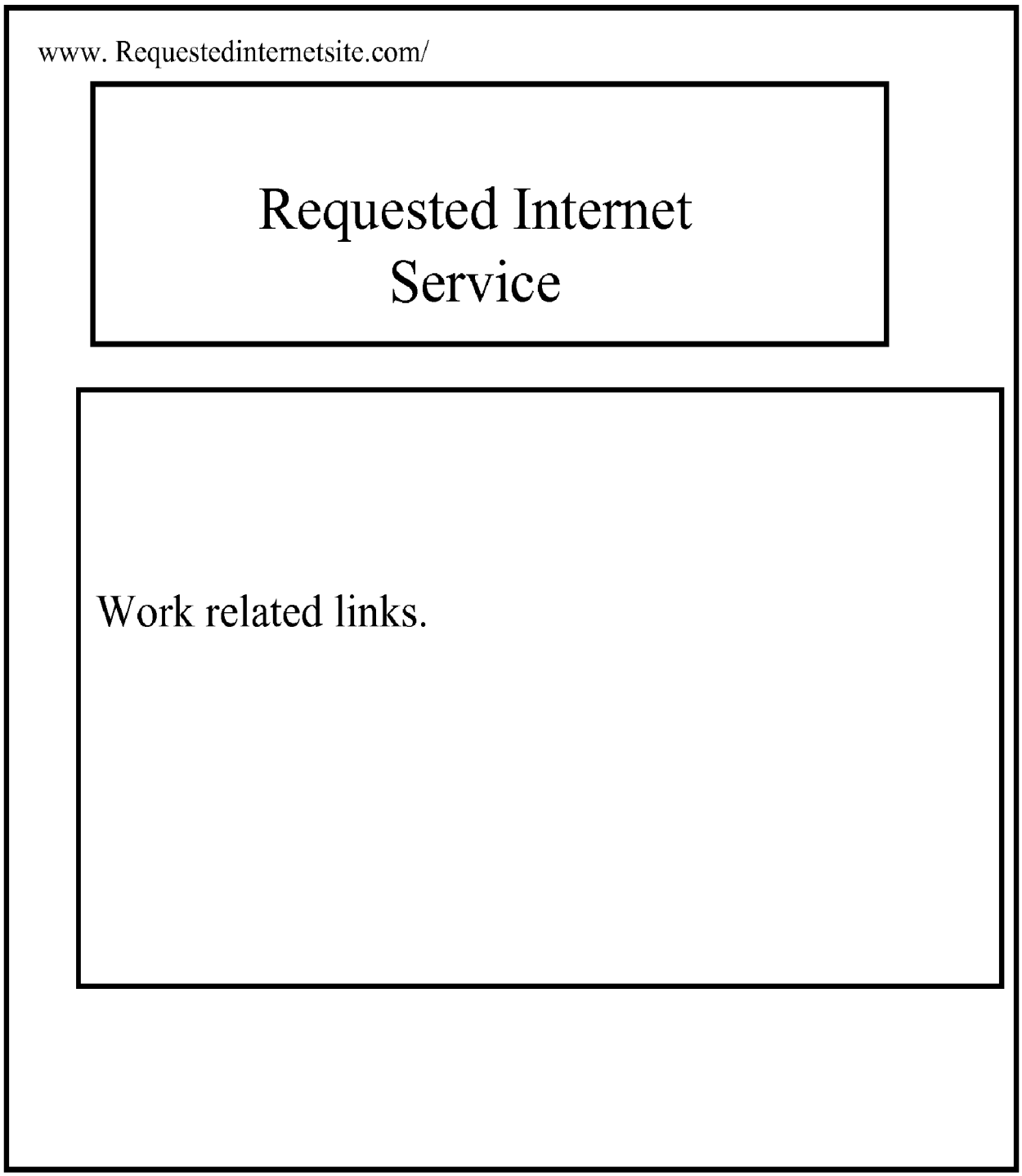

www. Requestedinternetsite.com/
Requested Internet Service
Work related links.
400

FIG. 4

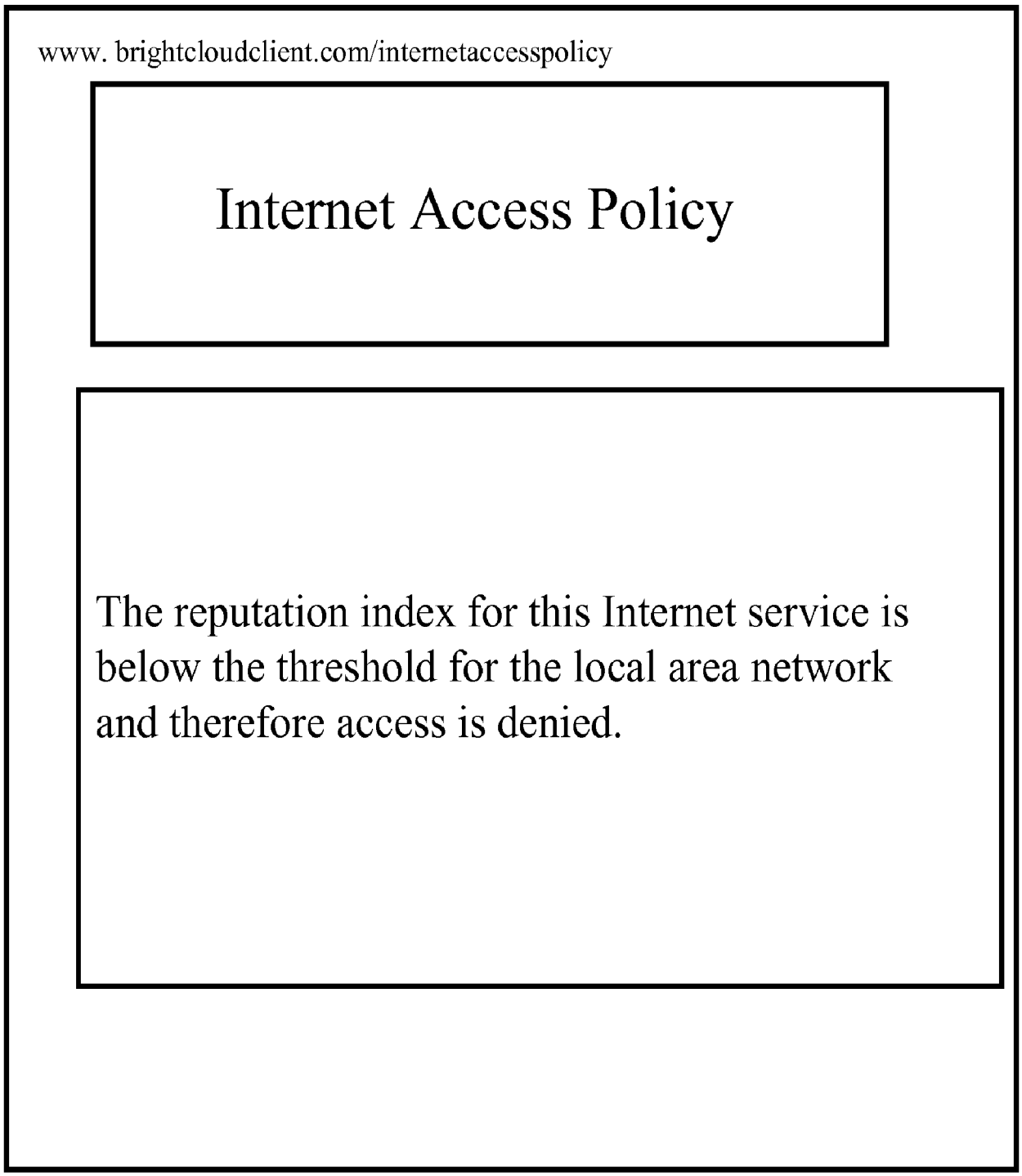
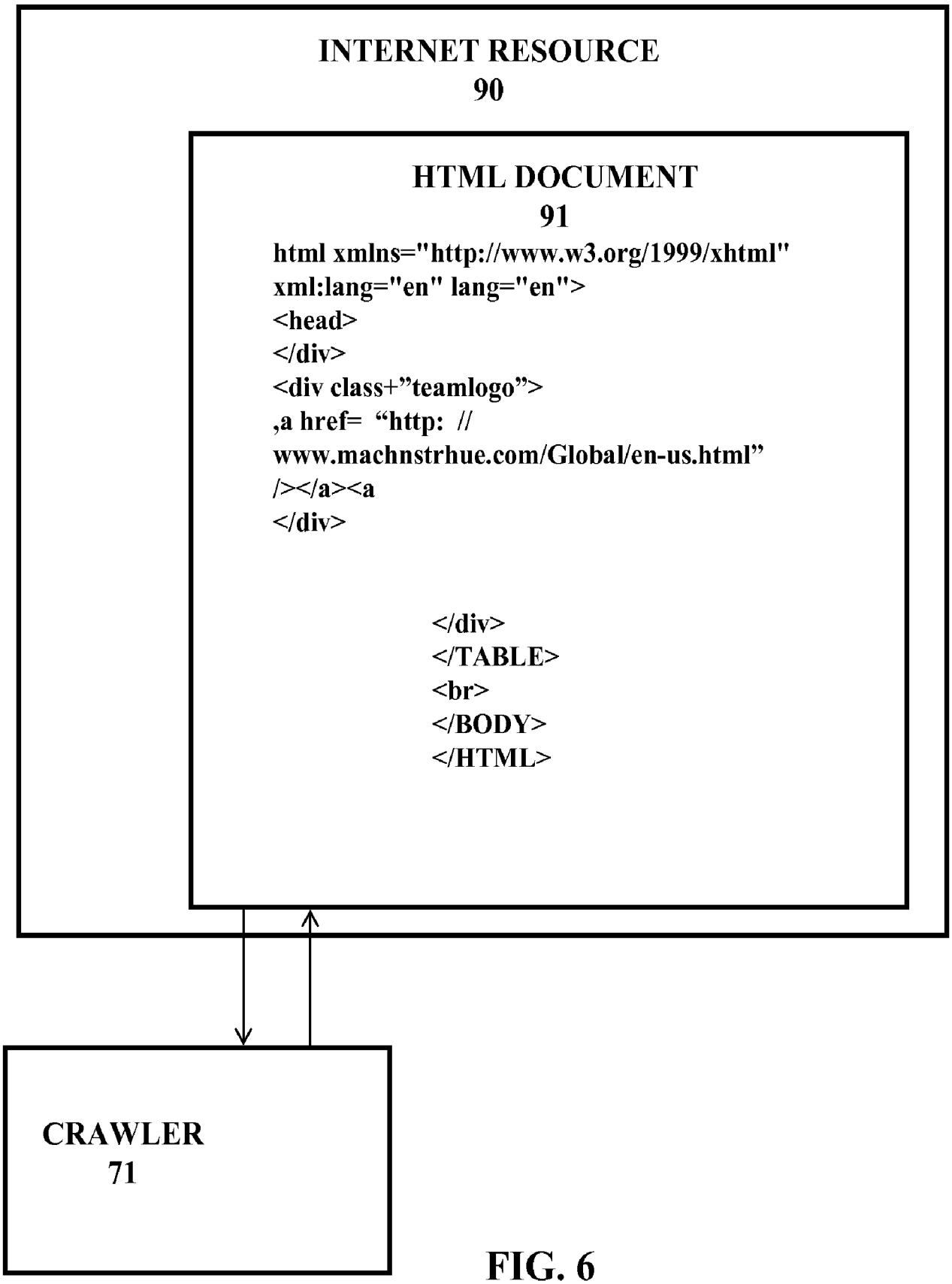

www. brightcloudclient.com/internetaccesspolicy
Internet Access Policy
The reputation index for this Internet service is below the threshold for the local area network and therefore access is denied.
500

FIG. 5


INTERNET RESOURCE
90
HTML DOCUMENT
91
html xmlns="http://www.w3.org/1999/xhtml"
xml:lang="en" lang="en">
<head>

<div class+"teamlogo">
,a href= "http: //
www.machnstrhue.com/Global/en-us.html"
/><a

</TABLE>

</BODY>
</HTML>
CRAWLER
71

FIG. 6

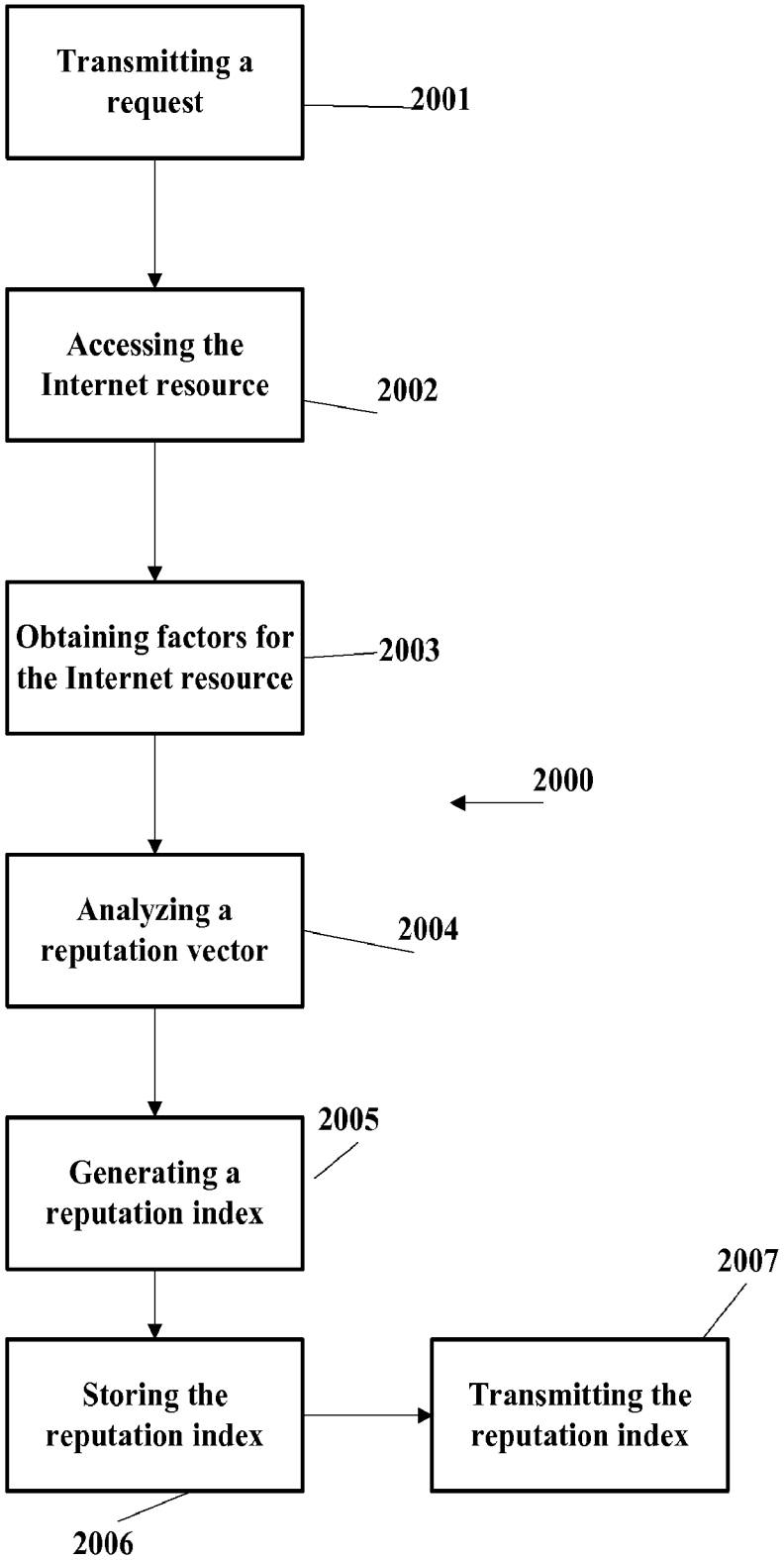

Transmitting a request
2001
Accessing the Internet resource
2002
Obtaining factors for the Internet resource
2003
2000
Analyzing a reputation vector
2004
2005
Generating a reputation index
2007
Storing the reputation index
Transmitting the reputation index
2006

FIG. 7

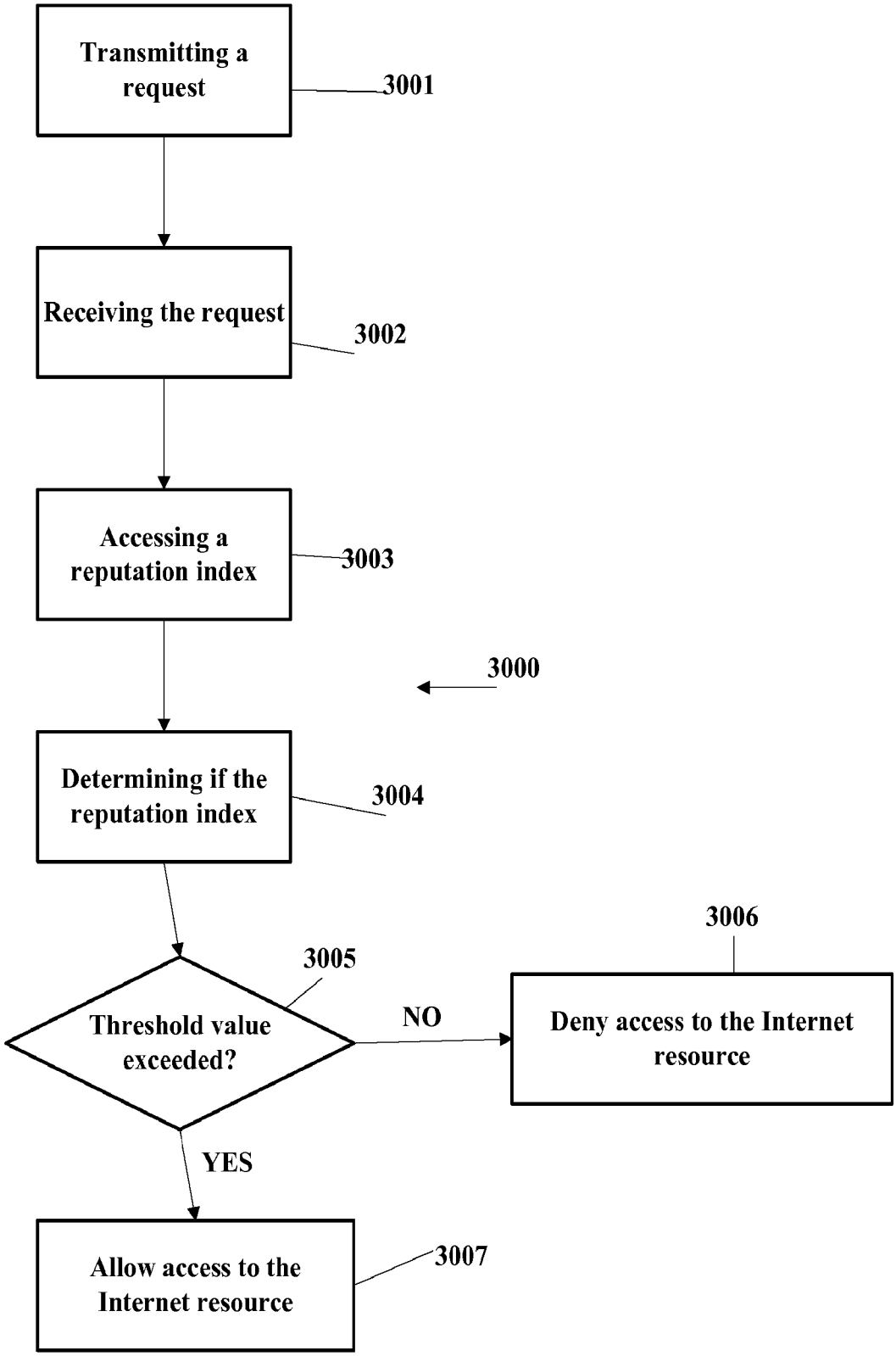

Transmitting a request
3001
Receiving the request
3002
Accessing a reputation index
3003
3000
Determining if the reputation index
3004
3006
3005
Threshold value exceeded?
NO
Deny access to the Internet resource
YES
Allow access to the Internet resource
3007

FIG. 8

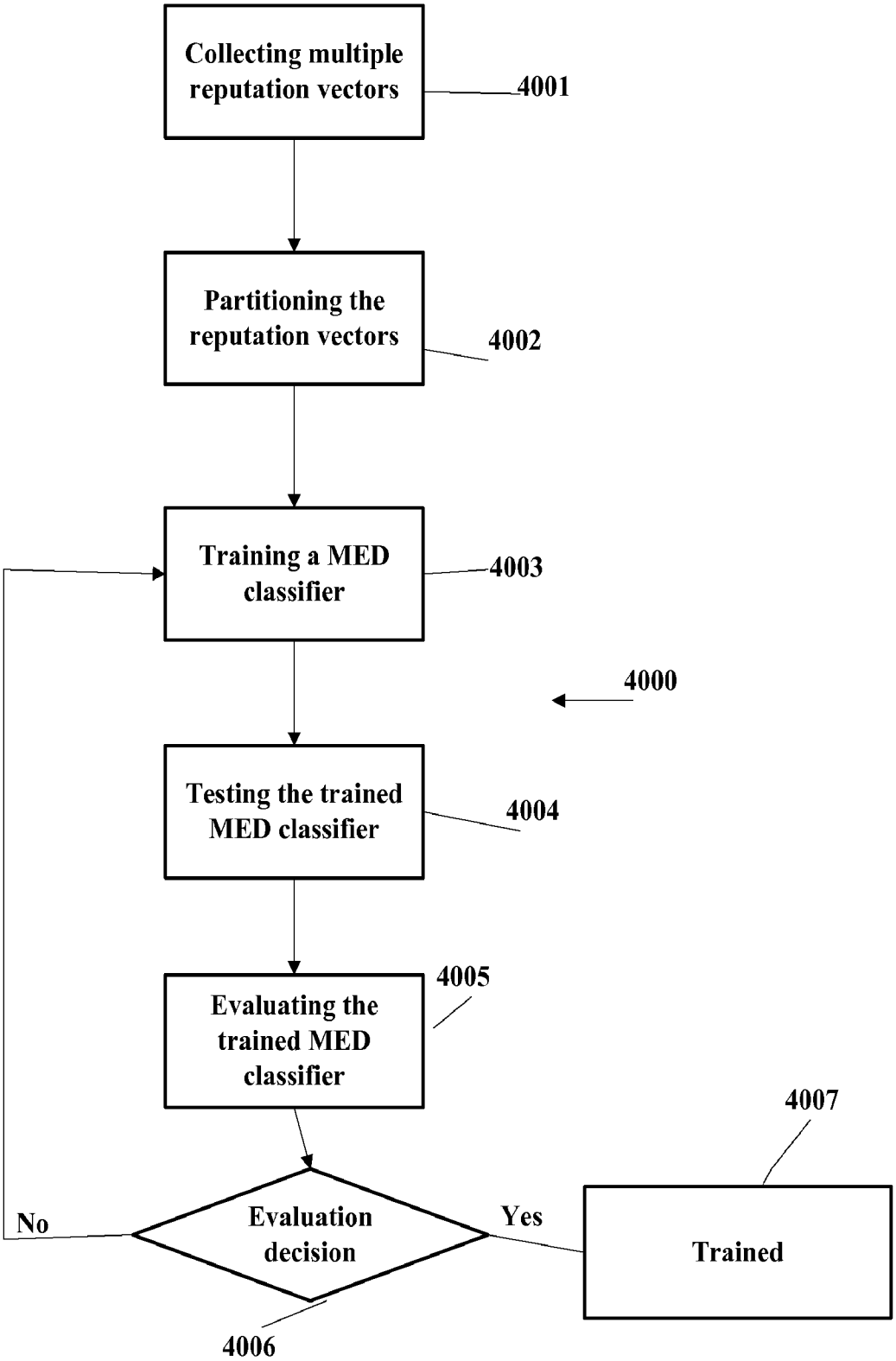

Collecting multiple reputation vectors
4001
Partitioning the reputation vectors
4002
Training a MED classifier
4003
4000
Testing the trained MED classifier
4004
Evaluating the trained MED classifier
4005
4007
Evaluation decision
No
Yes
Trained
4006

FIG. 9

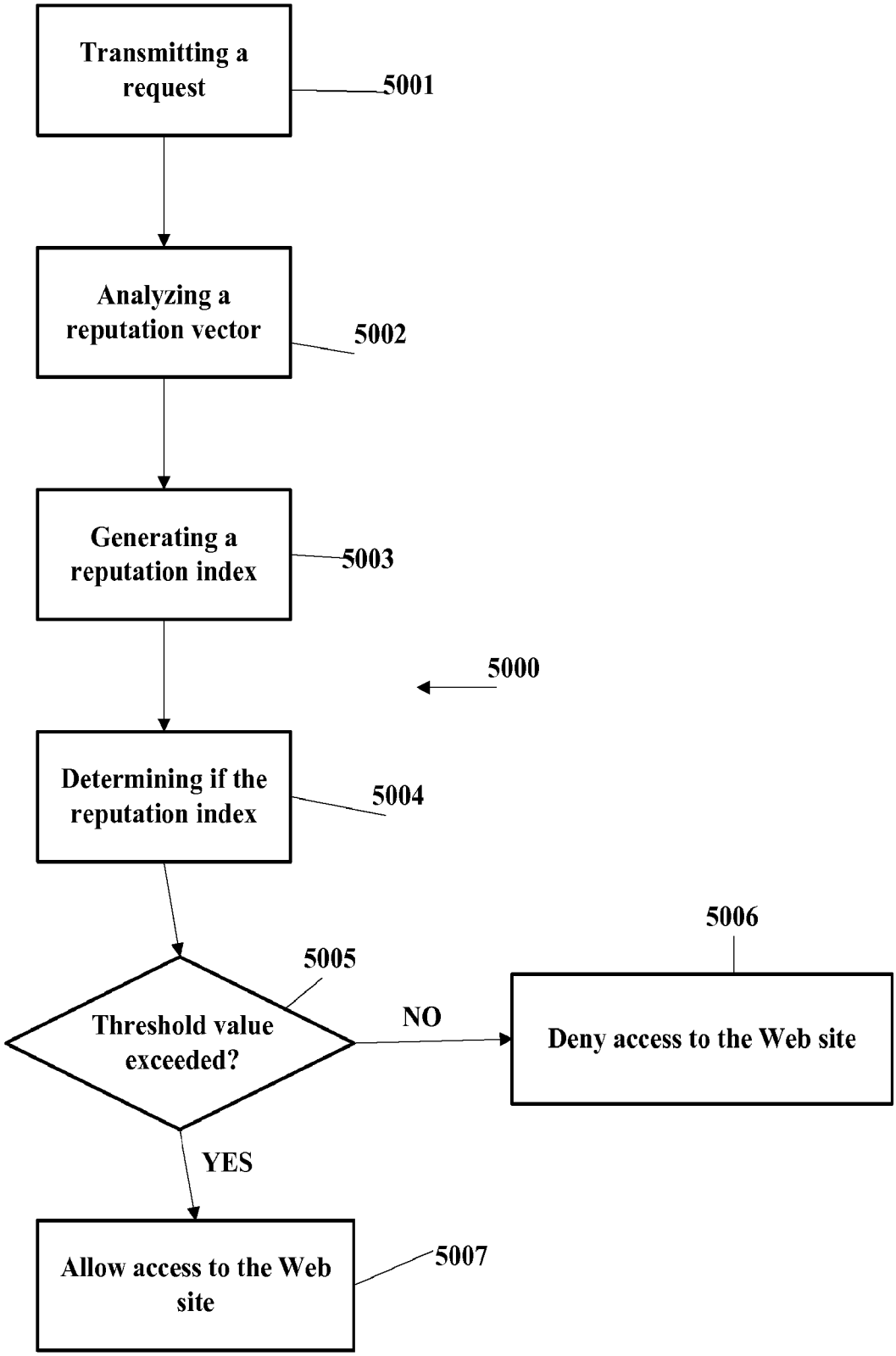

Transmitting a request
5001
Analyzing a reputation vector
5002
Generating a reputation index
5003
5000
Determining if the reputation index
5004
5006
5005
Threshold value exceeded?
NO
Deny access to the Web site
YES
Allow access to the Web site
5007

FIG. 10

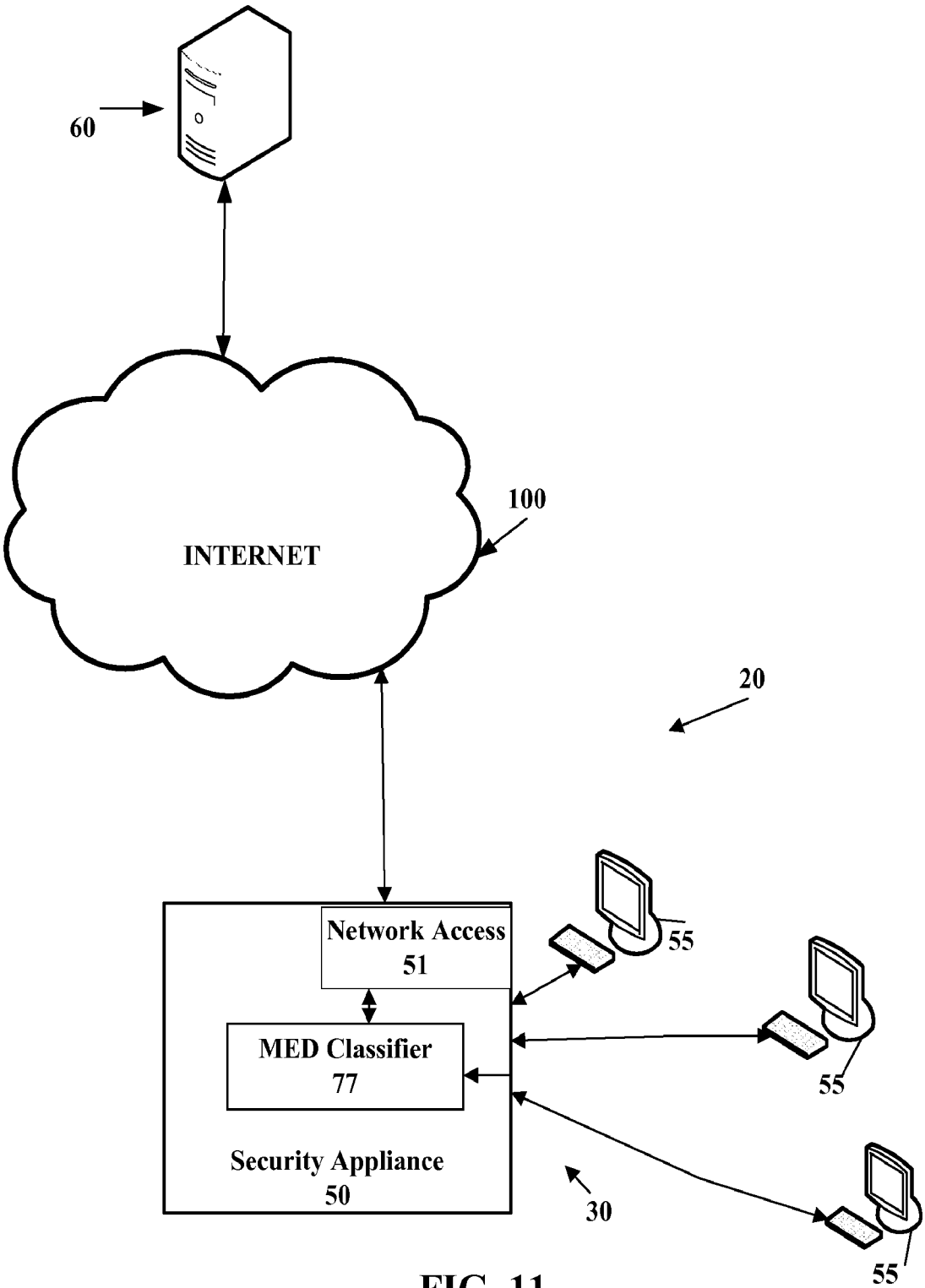

60
INTERNET
100
20
Network Access
51
MED Classifier
77
Security Appliance
50
30
55
55
55

FIG. 11