PROXY APPARATUS, ANALYSIS APPARATUS, REFERENCE APPARATUS, AND METHOD, PROGRAM, AND SYSTEM EXECUTED THEREBY

BACKGROUND

1. TECHNICAL FIELD

[0001] The present invention relates to a proxy apparatus, an analysis apparatus, a reference apparatus, a security risk determination method, an analysis method, a reference method, a security risk determination program, an analysis program, a reference program, and a security risk determination system.

2. RELATED ART

[0002] There is known a using service identifying apparatus which activates an Internet browser to access an analysis target website, monitors an HTTP request to extract a URL of an external request, and refers to an external service database to identify a service corresponding to a domain and a path of the URL of the external request (see, for example, Patent Document 1). However, in such a method, when a reference destination of the external request is malicious or invalid, there is a risk that the using service identifying apparatus is infected with malware or the like.

Related Art Document

Patent Document

Patent Document 1: Japanese Patent Application Publication No. 2018-185741

General Disclosure

[0003] (Item 1)

A proxy apparatus may include a request determination unit which receives a request transmitted from a virtual browser and determines a security risk of a request URL specified in the request. The proxy apparatus may include a request transfer unit which transfers, to a server operating an analysis target site, the request for which the request determination unit determines that the request URL is normal, the analysis target site being a target site for which a service used

at the analysis target site is identified. The proxy apparatus may include a response analysis unit which receives and analyzes a response from the server. The proxy apparatus may include a response return unit which returns a response to the virtual browser.

(Item 2)

The response analysis unit may determine a security risk from the response transmitted from the server. The response return unit may return, to the virtual browser, the response determined by the response analysis unit to have no security risk.

(Item 3)

The proxy apparatus may further include a request blocking unit which blocks the request when the request determination unit determines that the request URL is abnormal.

(Item 4)

When it is determined that the request URL is abnormal, the request blocking unit may generate a new response indicating that the request URL has a security risk and return the new response to the virtual browser.

(Item 5)

The proxy apparatus may further include a response blocking unit which blocks return of the response when the response analysis unit determines that the response has a security risk.

(Item 6)

When the response analysis unit determines that the response has a security risk, the response blocking unit may further generate a new response indicating that a page included in the response has a security risk and return the new response to the virtual browser.

(Item 7)

When a status code of the response transmitted from the server is not 200, the response analysis unit may determine a page included in the response as a non-referable page.

(Item 8)

The virtual browser may be activated by an analysis apparatus in order to specify a site

URL of the analysis target site, transmit the request to a server operating the analysis target site, and receive the response from the server.

[0004] (Item 9)

An analysis apparatus may include an analysis target acquisition unit which acquires a site URL of an analysis target site. The analysis apparatus may include an analysis execution unit which analyzes one or more pages in a site published at the site URL. The analysis execution unit may include a site access unit which specifies the site URL, transmits the request to the proxy apparatus according to item 1, and receives the response from the proxy apparatus by using the virtual browser. The analysis execution unit may include a status extraction unit which extracts a security risk from the response.

[0005] (Item 10)

The analysis execution unit may further include a link extraction unit which extracts a page URL of a linked page from the response. The site access unit may further specify a page URL of the linked page, transmit a request to the proxy apparatus, and receive a response from the proxy apparatus.

(Item 11)

The analysis apparatus may further include a status registration unit which registers, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

[0006] (Item 12)

A reference apparatus may include a status acquisition unit which acquires page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by the analysis apparatus

according to item 11. The reference apparatus may include a status output unit which outputs, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

[0007] (Item 13)

A security risk determination method executed by a computer may include receiving a request transmitted from a virtual browser and determining a security risk of a request URL specified in the request. The security risk determination method may include requesting by transferring, to a server operating an analysis target site, the request for which it is determined in the determining that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified. The security risk determination method may include receiving and analyzing a response from the server. The security risk determination method may include returning a response to the virtual browser.

[0008] (Item 14)

An analysis method executed by a computer may include acquiring a site URL of an analysis target site. The analysis method may include analyzing one or more pages in a site published at the site URL. The analyzing may include specifying the site URL, transmitting the request to a proxy apparatus including a computer which executes the security risk determination method according to item 13, and receiving the response from the proxy apparatus by using the virtual browser. The analyzing may include extracting a security risk from the response.

(Item 15)

The analysis method may further include registering, in a status database, each of page URLs of the one or more pages and a status code and a security risk of the response to the request specifying each of the page URLs in association with one another.

[0009] (Item 16)

A reference method executed by a computer may include acquiring page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by the analysis method according to item 15. The reference method may include outputting, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

[0010] (Item 17)

A security risk determination program may cause a computer to execute receiving a request transmitted from a virtual browser and determining a security risk of a request URL specified in the request. The security risk determination program may cause the computer to execute transferring, to a server operating an analysis target site, the request for which it is determined in the determining that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified. The security risk determination program may cause the computer to execute receiving and analyzing a response from the server. The security risk determination program may cause the computer to execute returning a response to the virtual browser.

[0011] (Item 18)

An analysis program may cause a computer to execute acquiring a site URL of an analysis target site. The analysis program may cause a computer to execute analyzing one or more pages in a site published at the site URL. The analyzing may include specifying the site URL, transmitting the request to a proxy apparatus including a computer which executes the security risk determination program according to item 17, and receiving the response from the proxy apparatus by using the virtual browser. The analyzing may include extracting a security risk from the response.

(Item 19)

The analysis program may cause the computer to further execute registering, in a status

database, each of page URLs of the one or more pages and a status code and a security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 20)

A reference program may cause a computer to execute acquiring page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by executing the analysis program according to item 19. The reference program may cause the computer to execute outputting, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

[0012] (Item 21)

A security risk determination system may include the proxy apparatus according to item 1. The security risk determination system may include the analysis apparatus according to item 9. The security risk determination system may include the reference apparatus according to item 12. The security risk determination system may include a status database in which data in which page URLs of one or more pages are associated with a status code and a security risk of a response to a request specifying each of the page URLs is registered.

[0013] (Item 22)

The security risk determination system may include an analysis apparatus which identifies a service used at an analysis target site. The security risk determination system may include a proxy apparatus which mediates a server operating the analysis target site and a virtual browser activated by the analysis apparatus. The analysis target site may be a target site for which a service used at the analysis target site is identified. The analysis apparatus may execute acquiring a site URL of the analysis target site, and specifying the site URL and transmitting a request to the proxy apparatus by using a virtual browser. The proxy apparatus may execute receiving the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request. The proxy apparatus may execute transferring the request for which it is

determined that the request URL is normal, to a server operating an analysis target site. The proxy apparatus may execute receiving and analyzing a response from the server. The proxy apparatus may execute returning a response to the virtual browser. The analysis apparatus may execute receiving the response from the proxy apparatus, extracting a security risk from the response, analyzing one or more pages in a site published at the site URL based on the response, and registering, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 23)

When it is determined that the request URL is abnormal, the proxy apparatus may generate a new response indicating that the request URL has a security risk and return the new response to the virtual browser.

(Item 24)

The proxy apparatus may further execute blocking communication with the analysis target site when it is determined that the response has a security risk.

(Item 25)

In the blocking, when it is determined that the response has a security risk, the proxy apparatus may further generate a new response indicating that a page included in the response has a security risk and return the new response to the virtual browser.

(Item 26)

A reference apparatus may further execute outputting the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs to a user terminal from the status database registered in the registering.

[0014] (Item 27)

A security risk determination method executed by one or more apparatuses may include

by an analysis apparatus, acquiring a site URL of an analysis target site, and specifying the site URL and transmitting a request to a proxy apparatus by using a virtual browser, the analysis target site being a target site for which a service used at the analysis target site is identified. The security risk determination method may include receiving, by the proxy apparatus, the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request. The security risk determination method may include transferring, by the proxy apparatus, the request for which it is determined that the request URL is normal, to a server operating an analysis target site. The security risk determination method may include receiving and analyzing, by the proxy apparatus, a response from the server. The security risk determination method may include returning, by the proxy apparatus, a response to the virtual browser. The security risk determination method may include, by the analysis apparatus, receiving the response from the proxy apparatus, extracting a security risk from the response, analyzing one or more pages in a site published at the site URL based on the response, and registering, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

[0015] (Item 28)

A security risk determination program may cause one or more computers to execute acquiring a site URL of an analysis target site, and specifying the site URL and transmitting a request to a proxy apparatus by using a virtual browser, the analysis target site being a target site for which a service used at the analysis target site is identified. The security risk determination program may cause the one or more computers to execute receiving, by the proxy apparatus, the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request. The security risk determination program may cause the one or more computers to execute transferring, by the proxy apparatus, the request for which it is determined that the

request URL is normal, to a server operating an analysis target site. The security risk determination program may cause the one or more computers to execute receiving and analyzing, by the proxy apparatus, a response from the server. The security risk determination program may cause the one or more computers to execute returning, by the proxy apparatus, a response to the virtual browser. The security risk determination program may cause the one or more computers to execute receiving the response from the proxy apparatus, extracting a security risk from the response, analyzing one or more pages in a site published at the site URL based on the response, and registering, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

[0016] The summary clause does not necessarily describe all necessary features of the embodiments of the present invention. In addition, the present invention may also be a sub-combination of the features described above.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] Fig. 1 illustrates a schematic configuration of a using service identifying system according to a first embodiment.

Fig. 2 illustrates a functional configuration of an analysis apparatus.

Fig. 3 illustrates an example of registration information of a site database.

Fig. 4 illustrates an example of data included in an HTTP response transmitted from an analysis target server.

Fig. 5 illustrates an example of registration information of an analysis item database.

Fig. 6 illustrates an analysis flow.

Fig. 7 illustrates a functional configuration of a using service identifying apparatus.

Fig. 8A illustrates an example of registration information of a service database.

Fig. 8B illustrates an example of the registration information of the service database.

Fig. 9 illustrates an example of registration information of a using service database.

Fig. 10 illustrates a using service identifying flow.

Fig. 11 illustrates a functional configuration of a reference apparatus.

Fig. 12 illustrates a screen for displaying a cookie, web storage, or an external communication URL corresponding to a using service acquired by an analysis item acquisition unit.

Fig. 13 illustrates a screen for displaying a site URL, a cookie in a site, web storage, or an external communication URL acquired by the analysis item acquisition unit.

Fig. 14A illustrates a using service reference flow.

Fig. 14B illustrates an analysis target input screen.

Fig. 15 illustrates an analysis item reference flow.

Fig. 16 illustrates a schematic configuration of a using service identifying system according to a second embodiment.

Fig. 17 illustrates a functional configuration of a proxy apparatus.

Fig. 18 illustrates an example of data included in an error response.

Fig. 19 illustrates a security risk determination flow.

Fig. 20 illustrates a functional configuration of the analysis apparatus.

Fig. 21 illustrates an example of registration information of a status database.

Fig. 22 illustrates an analysis flow.

Fig. 23 illustrates a functional configuration of the reference apparatus.

Fig. 24 illustrates a screen for displaying status information.

Fig. 25 illustrates a reference flow.

Fig. 26 illustrates an example of a configuration of a computer.

DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0018] (First Embodiment)

Hereinafter, the present invention will be described through embodiments of the present invention, but the following embodiments do not limit the scope of the present invention according to claims. In addition, not all of the combinations of features described in the embodiments are essential to the solution of the invention.

[0019] Fig. 1 illustrates a schematic configuration of a using service identifying system (also simply referred to as a system) 1 of a first embodiment. The system 1 is a network system which analyzes a website (also simply referred to as a site) to identify a using service used at the site, and provides information regarding the identified using service to a user terminal 90 or the like used by an administrator of the website. The system 1 includes an analysis apparatus 10, a using service identifying apparatus 50, and a reference apparatus 80. The system 1 may further include a site database (site DB) 20, an analysis item database (analysis item DB) 40, a service database (service DB) 60, and a using service database (using service DB) 70. These components are communicably connected to one another via a network 99. Note that the database is also referred to as a DB.

[0020] Note that the network 99 is a communication network that connects the analysis apparatus 10, the using service identifying apparatus 50, the reference apparatus 80, and various databases such as the site database 20 so that they can communicate with one another. The network 99 is, as an example, the Internet, but is not limited thereto, and any communication network such as a local area network or a telephone line may be adopted as long as these apparatuses and DBs can communicate with one another, or a network in which a plurality of communication networks are mixed may be adopted.

[0021] Note that an analysis target server 30 is a server which operates an analysis target site

that is a target for identifying a using service. Not only one analysis target site but also a plurality of analysis target sites may be analyzed, and in such a case, the system 1 accesses a plurality of analysis target servers 30 via the network 99. In the present example, only one analysis target server 30 is illustrated for simplicity.

[0022] Note that the user terminal 90 is, for example, a terminal apparatus used by an administrator of a website operated by the analysis target server 30. The terminal apparatus may be a computer terminal such as a personal computer, or a portable terminal such as a smartphone or a tablet computer. The user terminal 90 includes a display 91 such as a liquid crystal display, and functions as a client terminal that communicates with the reference apparatus 80, acquires information regarding a using service or the like used at a website, and causes the display 91 to display the information. A plurality of user terminals 90 respectively held by a plurality of users may be included in the system 1, and the plurality of users may be allowed to use the system 1.

[0023] The analysis apparatus 10 is a computer apparatus which reads a page in a site to be analyzed, transmits an HTTP/HTTPS request (also simply referred to as a request) to the analysis target server 30 operating the site by using a virtual browser, analyzes a received HTTP/HTTPS response (also simply referred to as a response) to extract a cookie or an external communication URL, acquires data stored in web storage of the virtual browser, and registers acquisition results thereof as analysis items. Note that the analysis items can be used by the using service identifying apparatus 50 to be described later. The analysis apparatus 10 includes a central processing unit (CPU) and a communication apparatus (both not illustrated). Note that the analysis apparatus 10 may be implemented by a cloud (a plurality of distributed servers, a plurality of subsystems, or the like). The CPU executes an analysis program to exhibit a function of analyzing one or more pages in a site to be analyzed and extracting a cookie issued by each page, web storage set by each page, and an external communication URL of external communication generated by each page. A

functional configuration of the analysis apparatus 10 will be described later. Note that the analysis program is activated, for example, by being stored in a ROM (not illustrated) and read by the CPU or being stored in a storage medium such as a CD-ROM and read by the CPU using a reading apparatus (not illustrated) to be developed in a RAM. The communication apparatus is an apparatus capable of performing wired or wireless communication with the using service identifying apparatus 50, the reference apparatus 80, and various databases via the network 99, and can perform communication using a protocol such as TCP/IP as an example.

[0024] Note that the virtual browser is a browser which operates on one or more independent virtual machines or on a container that shares a kernel of a host OS. All functions of commonly used browsers can also be used in the virtual browser, and all functions described for the browser in the present specification also apply to the virtual browser.

[0025] Fig. 2 illustrates a functional configuration of the analysis apparatus 10. The analysis apparatus 10 includes an analysis target acquisition unit 11, an analysis execution unit 12, and an analysis item registration unit 13.

[0026] The analysis target acquisition unit 11 is a functional unit which acquires a site URL of an analysis target site from the site database 20. Note that a method by which the analysis target acquisition unit 11 acquires the site URL is not limited to this, and may be acquired, for example, by an operator of the system 1 inputting the site URL via an input device (not illustrated).

[0027] The site database 20 is a storage apparatus in which various kinds of information such as a site URL of at least one analysis target site operated by the analysis target server 30 are registered (stored). A hard disk drive (HDD) or the like may be included as the storage apparatus.

[0028] Fig. 3 illustrates an example of information (URL information) regarding the site URL of the analysis target site registered in the site database 20. In the present example, three site URLs such as "XXX.com" are registered as the analysis target site. Here, the URL generally comprises a

scheme, a domain, a path, a parameter, and a fragment. A format of the site URL to be registered in the site database 20 is not particularly limited, and for example, the path and subsequent parts (a path, a parameter, and a fragment) of the URL may be expressed by a regular expression. These pieces of URL information may be registered by the operator of the system 1 via the input device (not illustrated).

[0029] The analysis execution unit 12 is a functional unit which analyzes one or more pages in a site published at a site URL for an analysis target site (for example, "XXX.com") acquired from the site database 20 by the analysis target acquisition unit 11, and extracts cookies issued by the one or more pages, web storage set by the one or more pages, or external communication URLs of external communication generated by the one or more pages. The analysis execution unit 12 includes a site access unit 121, an analysis item extraction unit 122, a script execution unit 123, and a link extraction unit 124.

[0030] Here, the cookie is data stored by a browser, and there are two types of cookies: a persistent cookie and a session cookie. The persistent cookie is stored for a specified period even when the browser is closed in a storage of a device on which the browser operates, and is used to retain information such as contents of a shopping cart of a user, language setting, and the like. The session cookie is a temporary cookie that is automatically deleted when the browser is closed and exists only during a session period, and is used to retain login information and the like when the user moves between pages. The cookie is extracted by analyzing a response header and acquiring cookie storage of the browser. The cookie has a Name attribute, such as a userPref attribute and a sessionId attribute, which is an identifier, a Domain attribute (domain attribute) for managing a domain in which a cookie is valid, a Path attribute (path attribute) for managing a path in which a cookie is valid, an Expires/Max-Age attribute for managing an expiration time (discard time) of a cookie, a Secure attribute for managing whether or not a cookie is transmitted only in

HTTPS communication, an HttpOnly attribute for managing whether or not to prevent access by JAVASCRIPT (registered trademark), a SameSite attribute for managing whether or not to permit transmission of a cross-site request, and the like.

[0031] A cookie with no specified domain attribute is valid only for a set Origin. The Origin comprises a domain, a protocol such as HTTP or HTTPS, and a port number. For example, a cookie set in "sub.example.com" with no specified domain attribute is valid only for "sub.example.com" and is invalid for "example.com" and other subdomains.

[0032] A cookie with a specified domain attribute is valid for the specified domain and its subdomains. For example, a cookie with a domain attribute specified as "example.com" is valid for "example.com" and its subdomains and is transmitted with all requests to the specified domain and its subdomains.

[0033] Note that a cookie in which no domain attribute is specified or in which a same domain as that of the accessing site is specified is referred to as a first-party cookie, and a cookie in which a different domain (external domain) is specified is referred to as a third-party cookie.

[0034] The web storage is an API for storing data in a browser, and there are two types of storage: local storage and session storage. The local storage is an API for storing data until the data is explicitly deleted in the browser. The data is stored in a data storage area of the browser and can be retained even when the browser is closed. The data stored in the local storage is shared within a same origin. The session storage is an API for storing data in a browser for only one session. The data is stored in a memory area reserved by the browser and is erased when the browser tab is closed. Note that the data stored in the web storage is also simply referred to as web storage.

[0035] The external communication URL is a URL specified as a communication destination of communication generated by a page, and belongs to a domain different from a domain of the page.

[0036] The site access unit 121 is a functional unit which specifies a site URL (a page URL of a top page of a site), transmits a request to the analysis target server 30 which operates an analysis target site, and receives a response from the analysis target server 30 by using the virtual browser.

The site access unit 121 further determines presence or absence of a link URL that has not yet been specified among link URLs extracted by the link extraction unit 124, and when there is a link URL that has not yet been specified and accessed, the site access unit 121 specifies the link URL and transmits a request to the analysis target server 30, and receives a response from the analysis target server 30. The site access unit 121 can recursively analyze all pages traced from the top page by specifying the link URL, transmitting the request to the analysis target server 30, and receiving the response from the analysis target server 30.

[0037] The analysis item extraction unit 122 is a functional unit which extracts, from the response received by the site access unit 121, cookies issued by one or more pages or external communication URLs of external communication generated by the one or more pages, and further acquires data stored in web storage set in the virtual browser.

[0038] Fig. 4 illustrates an example of data included in an HTTP response transmitted in response to transmission of a request to the analysis target server 30 (for example, the site URL "XXX.com" registered in the site database 20 of Fig. 3). The HTTP/HTTPS response includes a response line (status line), a response header, and a response body. In the present example, the response line includes information 35 regarding a version and a status code of a HTTP protocol, the version is HTTP/1.1, and the status code is 200, which indicates that the request is processed normally. The response header includes generation date and time of the response and cookie setting 31, and indicates that the response is generated at 10:18:14 on June 9, 2024 and the set cookie can be used under the "example.com" domain. The response body includes an HTML document, and a script tag 33 is embedded in the HTML document. A code of the script tag 33 transmits JSON data to https://example.com/api and displays the response on a console.

[0039] The analysis item extraction unit 122 is a functional unit which extracts, from a virtual browser, a cookie and web storage stored in the virtual browser and extracts, from a

communication log of the virtual browser, an external communication URL of external communication generated in a site. The analysis item extraction unit 122 may acquire these pieces of information from response data. In this case, for example, the analysis item extraction unit 122 extracts the cookie setting 31 "sessionId_app1=abc123; Domain=example.com;" from the response header of the data illustrated in Fig. 4 and external communication URL setting 32 "https://example.com/api" from the response body. Note that a method of extracting the information included in the cookie, the web storage, the communication log, or the response of the virtual browser is not particularly limited, and for example, the information can be extracted by calling an API such as JAVASCRIPT (registered trademark), Python, or JAVA (registered trademark) from an interface driver which operates the virtual browser.

[0040] The script execution unit 123 is a functional unit which, when a script is included in the response transmitted by the analysis target server 30, executes the script by the virtual browser. The script is generally described in an HTML document, and the HTML document is included in a response body of an HTTP response (see the script tag 33 in Fig. 3). An execution result of the script can be confirmed by, for example, a log or the like output from an interface driver which operates the virtual browser. As a result, it is possible to confirm whether or not execution of the script is successful.

[0041] The link extraction unit 124 is a functional unit which extracts a page URL of a linked page from the response transmitted by the analysis target server 30. The page URL is described in an HTML document, and the HTML document is included in a response body of an HTTP response (see Fig. 4). In the present example, a link URL of a link to "page1.html" in the page is extracted by extracting a link tag 34 from the response body. As a result, the site access unit 121 can access a linked page.

[0042] The analysis item registration unit 13 is a functional unit which, when the analysis item

extraction unit 122 extracts a cookie issued by a page in an analysis target site or an external communication URL of external communication generated by the page, or when data is acquired from a set web storage, associates a site URL with the cookie issued by the page in the site, the set web storage, or the generated external communication URL of external communication, and registers them in the analysis item database 40 so as to be acquirable by the using service identifying apparatus 50.

[0043] The analysis item database 40 is a storage apparatus in which a correspondence between a site URL, and a cookie issued by a page in the site, web storage set by the page, or an external communication URL generated from the page is registered (stored) by the analysis item registration unit 13. A hard disk drive (HDD) or the like may be included as the storage apparatus.

[0044] Fig. 5 illustrates an example of analysis information in which a site URL registered in the analysis item database 40 is associated with a cookie issued by a page in a site, web storage set by the page, or an external communication URL generated on the page. The analysis item database 40 includes three tables which associate a site URL with a cookie, the site URL with web storage, and the site URL with an external communication URL. Note that the analysis information may be managed by one table or may be managed by another method.

[0045] An upper table indicates cookie analysis information. The cookie analysis information includes a site URL column, a page URL column of a page that has issued the cookie, a Name attribute (cookie name) "Name" column of the cookie, a domain attribute "Domain" column, and a control classification "Control Type" column. In the present example, a page URL "XXX.com/" of a page of the site URL "XXX.com", a Name attribute "sessionId_app1" and a domain attribute "example.com" of a cookie issued by this page (see the cookie setting 31 in Fig. 4), and a control category "3rd Party Cookie" of this cookie are registered in association with one another. Note that in addition thereto, other attributes that can be possessed by the cookie may be registered in

the table.

[0046] A middle table indicates analysis information regarding the web storage. The analysis information of the web storage includes a site URL column, a page URL column of a page for which the web storage is set, an Origin "Origin" column, and a storage type "Storage Type" column. In the present example, a page URL "XXX.com/XXX" of a page of the site URL "XXX.com", an Origin "https://example1.com" of web storage set by the page, and a storage type "local Storage" are registered in association with one another. Note that in addition thereto, other information that can be included in the web storage, for example, a Key attribute and the like may be registered in the table.

[0047] A lower table indicates analysis information regarding the external communication URL. The analysis information regarding the external communication URL includes a site URL column, a page URL column of a page for which the web storage is set, and an external communication URL "External URL" column. In the present example, a page URL "XXX.com/YYY" of the page of the site URL "XXX.com" and an external communication URL "http://anotherdomain1.com" of external communication generated on this page are registered in association with one another. In addition, a page URL "XXX.com/ZZZ" of the page of the site URL "XXX.com" and an external communication URL "http://anotherdomain2.com" of external communication generated on this page are registered in association with one another.

[0048] The analysis apparatus 10 according to the present embodiment includes the analysis target acquisition unit 11 which acquires a site URL of an analysis target site, the analysis execution unit 12 which analyzes one or more pages in a site published at the site URL and extracts cookies issued by the one or more pages, web storage set by the one or more pages, or external communication URLs of external communication generated by the one or more pages, and the analysis item registration unit 13 which associates the site URL with the cookie, the web storage, or

the external communication URL issued by the page in the site and registers them so as to be acquirable by the using service identifying apparatus to be described later, so that the cookies issued by the one or more pages in the site published at the site URL, the web storage set by the one or more pages, or the external communication URLs of the external communication generated by the one or more pages can be extracted, and these data can be registered for each site so as to be acquirable by the using service identifying apparatus 50. It is possible not only to list and register the using services used at the analysis target site, but also to analyze and examine which of the cookie, the web storage, or the external transmission is used to access the using service.

[0049] Fig. 6 illustrates a flow of an analysis method executed by the analysis apparatus 10.

[0050] In step S11, the analysis apparatus 10 (analysis target acquisition unit 11) acquires a site URL of an analysis target site from the site database 20. There may be a plurality of analysis target sites. As described above, information (URL information) regarding the site URL of the analysis target site is registered in the site database 20. The URL information may be acquired by the operator of the system 1 inputting the URL information via the input device (not illustrated).

[0051] In step S12, the analysis apparatus 10 (analysis execution unit 12) analyzes a page in a site published at the site URL, and extracts all of cookies issued by the page, web storage set by the page, or external communication URLs of external communication generated by the page. In addition, all link URLs are extracted. First, the site access unit 121 specifies a site URL (a page URL of a top page of the site) or a page URL of a linked page extracted by the link extraction unit 124, transmits a request to the analysis target server 30 which operates the analysis target site, and receives a response from the analysis target server 30 by using the virtual browser. Next, the analysis item extraction unit 122 extracts, from the response received by the site access unit 121, cookies issued by one or more pages or external communication URLs of external communication generated by the one or more pages, and further acquires data stored in web storage set in the

virtual browser. Thereafter, when a script is included in the response transmitted by the analysis target server 30, the script execution unit 123 executes the script by the virtual browser. Then, the link extraction unit 124 extracts the page URL of the linked page from the response transmitted by the analysis target server 30.

[0052] As a result, for example, cookies with the Name attribute "sessionId_app1" and the domain attribute "example.com" issued by the page of the page URL "XXX.com/" of the analysis target site are extracted (see the cookie setting 31 in Fig. 4). Details of the extraction are as described above.

[0053] In step S13, the analysis apparatus 10 (site access unit 121) determines presence or absence of a link URL that has not yet been specified, among the link URLs extracted by the link extraction unit 124. If there is a link URL that has not yet been specified, the process returns to step S12, and if not, the process proceeds to step S14.

[0054] In step S14, the analysis apparatus 10 (site access unit 121) determines presence or absence of a site URL that has not yet been specified, among the site URLs extracted by the analysis target acquisition unit 11. If there is a site URL that has not yet been specified, for example, if analysis of "XXX.com" (see Fig. 3) as the analysis target site has been completed and analysis of "YYY.com" and "ZZZ.com" (see Fig. 3) as the analysis target sites has not yet been completed, the process returns to step S12, and if there is not the site URL (if the analysis of all the analysis target sites has been completed), the process proceeds to step S15.

[0055] In step S15, the analysis apparatus 10 (analysis item registration unit 13) associates the site URL with the cookie, the web storage, or the external communication URL, and registers them in the analysis item database 40 (see Fig. 5) so as to be acquirable by the using service identifying apparatus 50 (acquirable by a using service identifying method to be described later). As a result, for example, the site URL "XXX.com/", and the cookie with the Name attribute "sessionId_app1"

and the domain attribute "example.com" (see the cookie setting 31 in Fig. 4) issued by the page in the site are registered in association with one another. In this manner, all analysis information is accumulated in the analysis item database 40. Details of the accumulation of the analysis information are as described above.

[0056] The analysis method according to the present embodiment is executed by a computer, and includes: acquiring a site URL of an analysis target site; analyzing one or more pages in a site published at the site URL and extracting cookies issued by the one or more pages, web storage set by the one or more pages, or external communication URLs of external communication generated by the one or more pages; and associating the site URL of the analysis target site with the cookie issued by the page in the analysis target site, the web storage set by the page, or the external communication URL generated by the page and registering them so as to be acquirable by the using service identifying method to be described later. As a result, it is possible not only to list and register the using services used at the analysis target site, but also to analyze and examine which means of the cookie, the web storage, or the like is used to access the using service.

[0057] The using service identifying apparatus 50 is a computer apparatus that identifies a service used by an analysis target site. The using service identifying apparatus 50 includes a CPU and a communication apparatus (both not illustrated). Note that the using service identifying apparatus 50 may be implemented by a cloud (a plurality of distributed servers, a plurality of subsystems, or the like). The CPU executes a using service identifying program to exhibit a function of identifying a service used by the analysis target site. A functional configuration of the using service identifying apparatus 50 will be described later. Note that the using service identifying program is activated, for example, by being stored in a ROM (not illustrated) and read by the CPU or being stored in a storage medium such as a CD-ROM, and read by the CPU using a reading apparatus (not illustrated) to be developed in a RAM. The communication apparatus is an

apparatus capable of performing wired or wireless communication with the analysis apparatus 10, the reference apparatus 80, and various databases via the network 99, and can perform communication using a protocol such as TCP/IP as an example.

[0058] Fig. 7 illustrates a functional configuration of the using service identifying apparatus 50. The using service identifying apparatus 50 includes an analysis item acquisition unit 51, a using service identifying unit 52, and a using service registration unit 53.

[0059] The analysis item acquisition unit 51 is a functional unit which acquires, from the analysis item database 40, an analysis item (correspondence information registered in each of one or more rows illustrated in Fig. 5) including a correspondence relationship between a page URL of an analysis target site and a cookie issued by at least one page in the analysis target site, web storage set by the at least one page, or an external communication URL generated on the page. The analysis item acquisition unit 51 acquires analysis items one by one from the analysis item database 40 and transmits the analysis items to the using service identifying unit 52. The analysis item acquisition unit 51 can use an analysis item analyzed by an apparatus (that is, the analysis apparatus 10) separate from the using service identifying apparatus 50.

[0060] The using service identifying unit 52 is a functional unit which matches a cookie, web storage, or an external communication URL included in each analysis item acquired by the analysis item acquisition unit 51 from the analysis item database 40 with the service database 60 and identifies a service corresponding to matched data as a using service used on the site. The using service identifying unit 52 includes an external communication URL matching unit 521, a cookie matching unit 522, and a web storage matching unit 523. Here, the service is an external service such as an advertisement service or an access analysis service which is executed, for example, by specifying an external communication URL and transmitting a request.

[0061] The service database 60 is a storage apparatus in which one or more pieces of service

information in which each of a cookie, web storage, and an external communication URL is associated with a service is registered (stored). A hard disk drive (HDD) or the like may be included as the storage apparatus. Associating each of the cookie, the web storage, and the external communication URL with the service means associating each of the cookie, the web storage, and the external communication URL with an ID, a name, or the like of the service that can uniquely identify the service. In addition, the service is not limited to be represented by its name (referred to as a service name), and may be represented by an ID. In the present embodiment, as an example, the service is represented by the service name.

[0062] Fig. 8A illustrates a part of an example of service information registered in the service database 60. The service database 60 includes three tables which associate a service name with a cookie, the service name with web storage, and the service name with an external communication URL.

[0063] An upper table indicates service information in which a cookie is associated with a service. The service information includes a service name column, a cookie Name attribute (cookie name) "Name" column, and a domain attribute "Domain" column. In the present example, a service name "InsightPulse" is associated with a cookie with a Name attribute "app1" used by a service indicated by the service name and a domain attribute "example.com". Note that in addition thereto, other attributes that can be possessed by the cookie may be registered in the table.

[0064] A middle table indicates service information in which web storage is associated with a service. The service information includes a service name column and an Origin "Origin" column of web storage. In the present example, a service name "MetricVision" is associated with an Origin "https://example1.com" of Web Storage used by a service indicated by the service name. Note that in addition thereto, other information that can be included in the web storage, for example, a Key attribute and the like may be registered in the table.

[0065] A lower table indicates service information in which an external communication URL is associated with a service. The service information includes a service name column and an external communication URL "External URL" column. In the present example, a service name "AnalyticaFlow" is associated with an external communication URL "anotherdomain2.com" specified for accessing the service indicated by the service name.

[0066] In addition to the service name, other service information may also be registered in the service database 60. The other service information is, for example, a type of service such as advertisement or access analysis, a service provider of the external service, or the like.

[0067] The external communication URL matching unit 521 is a functional unit which, when an external communication URL is included in an analysis item acquired from the analysis item database 40, extracts, from the service database 60, a service name corresponding to data at least partially matching with the external communication URL, and identifies this service name as a using service. For example, the external communication URL matching unit 521 acquires data of a second row having "anotherdomain2.com" in the ExternalURL column of the lower table from the analysis information registered in the analysis item database 40 (see Fig. 5), performs a partial match search on the ExternalURL column of the service information of the lower table registered in the service database 60 by using the "anotherdomain2.com" key (see Fig. 8A), and acquires, as the using service, a service name "AnalyticaFlow" of the service information in a first row having a value of the ExternalURL column matching with "anotherdomain2.com".

[0068] The cookie matching unit 522 is a functional unit which, when a cookie is included in the analysis item acquired from the analysis item database 40, extracts, from the service database 60, a service at least partially matching with a Name attribute of the cookie, and as necessary, extracts, from the service database 60, a service corresponding to data at least partially matching with the Name attribute of the cookie and completely matching with a domain attribute of the cookie, and

acquires the service as a using service. Here, "as necessary" means that, for example, when a type of cookie is a third-party cookie, processing is performed. In addition, when a plurality of pieces of data in the cookie data of the analysis item have at least partially matching Name attribute and completely matching domain attribute, the cookie matching unit 522 acquires, as the using service, a most frequent service among a plurality of services indicated by the plurality of pieces of data.

[0069] Fig. 8B illustrates a part of an example of the service information in which a cookie is associated with a service, registered in the service database 60. In the present example, the cookie matching unit 522 extracts three cookies specifying InsightPulse and one cookie specifying Service2, the cookies having a domain attribute completely matching with a domain attribute "example.com" of a cookie in a first row of the analysis item acquired from the upper table of the analysis item database 40 (see Fig. 5). Among these services, a most frequent service with a Name attribute matching with the Name attribute portion "app1" of the analysis item is InsightPulse, and thus the cookie matching unit 522 acquires InsightPulse as a using service. As a case where cookies having a matching domain attribute are distinguished by the Name attribute in this manner, for example, a case where InsightPulse manages a login state or the like of an entire platform and Service2 manages a sub-service in the platform, or the like can be assumed. The cookie matching unit 522 can identify a using service even when there is a plurality of services that can be access targets of a certain cookie.

[0070] The web storage matching unit 523 is a functional unit which, when web storage is included in the analysis item acquired from the analysis item database 40, extracts, from the service database 60, a service associated with data partially matching with an Origin of the web storage and acquires the service as a using service. For example, the web storage matching unit 523 acquires, from the middle table of the analysis item database 40, an analysis item in a first row,

extracts data in a first row of the middle table of the service database 60 which has an Origin partially matching (in the present example, completely matching) with the Origin "https://example1.com" included in the analysis item, and acquires a corresponding service "MetricVision" as a using service.

[0071] Note that the web storage matching unit 523 may extract a service associated with data having a completely matching Origin attribute. In addition, a service associated with data having a partially or completely matching Key attribute as well as a partially or completely matching Origin attribute may be extracted.

[0072] The using service registration unit 53 is a functional unit which, when the using service identifying unit 52 can identify a service name, registers a site URL and a using service in a using service database 70 in association with one another. The using service registration unit 53 registers, in the using service database 70, a cookie, web storage, or an external communication URL for which a service name cannot be identified, as data (for example, Unknown) not associated with the service. The using service registration unit 53 can accumulate information regarding the identified using service so that the information can be used by another apparatus. In addition, it is possible to detect a cookie or the like that is not associated with a service, that is, a cookie or the like that is not operating.

[0073] The using service database 70 is a storage apparatus in which a site URL and a service name of a using service used at a site thereof are registered in association with one another. A hard disk drive (HDD) or the like may be included as the storage apparatus.

[0074] Fig. 9 illustrates an example of using service information registered in the using service database 70 by the using service registration unit 53. The using service information is registered in a table format including a using service name column and a site URL (URL) column. In the present example, the site URL "XXX.com" described above and the service name "InsightPulse"

identified as a using service used at the site are registered in association with one another, the site URL "XXX.com" and the service name "MetricVision" identified as a using service used at the site are registered in association with one another, and the site URL "XXX.com" described above and the service name "AnalyticaFlow" identified as a using service used at the site are registered in association with one another. Since a service corresponding to an external communication URL used at the site URL "XXX.com" cannot be identified, the site URL "XXX.com" and "Unknown" which indicates that the using service is unknown are registered in association with one another.

[0075] The using service identifying apparatus 50 according to the present embodiment includes the analysis item acquisition unit 51 which acquires an analysis item including a correspondence relationship between a page URL of at least one page in an analysis target site and a cookie issued by the at least one page, web storage set by the at least one page, or an external communication URL of the external communication generated by the at least one page, the using service identifying unit 52 which matches the analysis item with a service database to identify, as a using service used at the analysis target site, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item, and the using service registration unit 53 which registers the site URL and the using service in association with one another, so that the cookie, the web storage, or the external communication URL for each analysis target site analyzed by the analysis apparatus 10 can be acquired, the corresponding using service can be identified, and the identified using service can be registered in the using service database 70.

[0076] Fig. 10 illustrates a flow of the using service identifying method executed by the using service identifying apparatus 50.

[0077] In step S21, the using service identifying apparatus 50 (analysis item acquisition unit 51) acquires, from the analysis item database 40 (see Fig. 5), all pieces of analysis information having a correspondence between a cookie, web storage, or an external communication URL and a site URL

of a site to which the analysis item belongs.

[0078] In step S22, the using service identifying apparatus 50 (using service identifying unit 52) matches the analysis item with the service database 60 (see Figs. 8A and 8B) to identify, as a using service used at an analysis target site, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item. First, the external communication URL matching unit 521 extracts, from the service database 60, a service corresponding to data at least partially matching with the external communication URL included in the analysis item acquired from the analysis item database 40. Next, the cookie matching unit 522 extracts, from the service database 60, a service corresponding to data at least partially matching with a Name attribute of a cookie included in each of the analysis items acquired from the analysis item database 40 and completely matching with a domain attribute of the cookie, and acquires the service as a using service. Here, when there is a plurality of pieces of data in the cookie data of the analysis item have at least partially matching Name attribute and completely matching domain attribute, the cookie matching unit 522 acquires, as the using service, a most frequent service among a plurality of services indicated by the plurality of pieces of data. Further, the web storage matching unit 523 extracts, from the service database 60, a service associated with data partially matching with an Origin attribute of web storage included in each of the analysis items acquired from the analysis item database 40, and acquires the service as a using service. Details are as described above. In this manner, it is preferable to perform matching with the service database 60 in order of matching of the external communication URL, matching of the cookie, and matching of the web storage. Since certainty of identifying the service is the highest for the external communication URL match, next for the cookie match, and lowest for the web storage match, performing the matching in this order enables efficient identification.

[0079] For example, for analysis information (URL column "XXX.com/ZZZ/" and ExternalURL

column "anotherdomain2.com") in a second row registered in the lower table of the analysis item database 40 (see Fig. 5), the service database 60 (see Fig. 8A) is searched with the external URL "anotherdomain2.com" as a key, and the service name "AnalyticaFlow" of the service information is acquired as a using service, based on information in a first row of the lower table. In addition, for analysis information (URL column "XXX.com/" and COOKIE column "app1; domain=example.com") in a first row registered in the upper table of the analysis item database 40 (see Fig. 5), the service database 60 (see Fig. 8B) is searched to extract three cookies specifying InsightPulse having a Domain attribute completely matching with "example.com" and an attribute name of the Name attribute partially matching with "app1", thereby acquiring InsightPulse as a using service. In addition, an analysis item in a first row of the middle table of the analysis item database 40 (see Fig. 5) is acquired, data in a first row of the middle table of the service database 60 (Fig. 8A) partially matching (in the present example, completely matching) with "https://example1.com" that is an Origin included in the analysis item is extracted, and a corresponding service "MetricVision" is acquired as a using service. Details of the acquisition of the using service are as described above.

[0080] In step S23, the using service identifying apparatus 50 (using service registration unit 53) registers, in the using service database 70 (see Fig. 9), using service information in which a site URL is associated with a using service. As a result, for example, the site URL "XXX.com" and the service name "InsightPulse" (see a first row in Fig. 9) identified as a using service used at the site are registered in association with one another, the site URL "XXX.com" and the service name "AnalyticaFlow" (see a fourth row in Fig. 9) identified as a using service used at the site are registered in association with one another, and the site URL "XXX.com" and the service name "MetricVision" (see a second row in Fig. 9) identified as a using service used at the site are registered in association with one another. Accordingly, the using service information can be

accumulated. Details of the accumulation of the using service information are as described above.

[0081] In step S24, the using service identifying apparatus 50 (using service registration unit 53) determines presence or absence of unregistered analysis information. If it is present, the process returns to step S21, and if it is not present, the process ends. As a result, all the analysis information can be registered.

[0082] The using service identifying method according to the present embodiment is a using service identifying method executed by a computer, the method including acquiring an analysis item including a correspondence relationship between a site URL in an analysis target site and a cookie, web storage, or an external communication URL issued by at least one page in the analysis target site, matching the analysis item with a service database to identify, as a using service used at the analysis target site, a service name corresponding to the cookie, the web storage, or the external communication URL included in the analysis item, and registering the site URL and the using service in association with one another, and in the service database, data which has, as a set, the service name and each of the cookie, the web storage, and the external communication URL is registered. According to this, it is possible to acquire a cookie, web storage, or an external communication URL for each site analyzed by the analysis apparatus 10, identify a corresponding using service, and register the identified using service in the using service database 70.

[0083] The reference apparatus 80 is a computer apparatus which outputs a service used at a target site in association with a cookie for accessing a using service, web storage, or an external communication URL. The reference apparatus 80 includes a CPU and a communication apparatus (both not illustrated). Note that the reference apparatus 80 may be implemented by a cloud (a plurality of distributed servers, a plurality of subsystems, or the like). The CPU executes a using service reference program to exhibit a function of outputting a service used at a target site. A

functional configuration of the reference apparatus 80 will be described later. Note that the using service reference program is activated, for example, by being stored in a ROM (not illustrated) and read by the CPU or being stored in a storage medium such as a CD-ROM and read by the CPU using a reading apparatus (not illustrated) to be developed in a RAM. The communication apparatus is an apparatus capable of performing wired or wireless communication with the analysis apparatus 10, the using service identifying apparatus 50, and various databases via the network 99, and can perform communication using a protocol such as TCP/IP as an example.

[0084] Fig. 11 illustrates a functional configuration of the reference apparatus 80. The reference apparatus 80 includes a using service reference unit 81 and an analysis item reference unit 82.

[0085] The using service reference unit 81 is a functional unit which outputs, to the user terminal 90, a service used by a site. Note that the "output" referred to here may be in a form of output to the display 91 of the user terminal 90, in a form of transmission in a file format to the user terminal 90, or in any other form. In addition to the using service, the using service reference unit 81 may output, to the user terminal 90, a cookie, web storage, or an external communication URL corresponding to the using service. The using service reference unit 81 includes a using service acquisition unit 811, a third analysis item acquisition unit 812, and a using service output unit 813.

[0086] The using service acquisition unit 811 is a functional unit which acquires, from the using service database 70, a using service registered by the using service identifying apparatus 50 as a service used at an analysis target site. For example, when "XXX.com" is specified as the analysis target site, the using service acquisition unit 811 acquires all rows including "XXX.com" in a URL column in using service information (see Fig. 9) registered in the using service database 70.

[0087] The analysis item acquisition unit (an example of the third analysis item acquisition unit)

812 is a functional unit which, for all using service information acquired by the using service acquisition unit 811, uses a using service name thereof as a key to acquire, from the service database 60, a cookie, web storage, or an external communication URL corresponding to the using service. For example, the analysis item acquisition unit 812 uses, as a key, the using service name "AnalyticaFlow" of the using service information acquired by the using service acquisition unit 811 to extract service information having a matching service name from the service database 60 illustrated in Fig. 8A, and acquires the external communication URL "anotherdomain2.com" included in the service information.

[0088] The using service output unit 813 is a functional unit which outputs, to the user terminal 90, the using service acquired by the using service acquisition unit 811. The using service output unit 813 further outputs a cookie, web storage, or an external communication URL corresponding to the using service acquired by the analysis item acquisition unit 812.

[0089] Fig. 12 illustrates an example of a display screen of the display 91 on which the cookie, the web storage, or the external communication URL corresponding to the using service acquired by the using service acquisition unit 811 is displayed. The screen includes a target site display unit 911 and a using service display unit 912. When the user specifies "XXX.com" as a reference target URL on the target site display unit 911, the reference apparatus 80 displays the site URL "XXX.com" on the target site display unit 911, and displays, on the using service display unit 912, the using service name used at the "XXX.com" acquired from the using service database 70 by the using service acquisition unit 811, and the cookie, the web storage, or the external communication URL acquired as ones corresponding to the using service from the service database 60 by the analysis item acquisition unit 812 in association with one another. As a result, for example, it is displayed that the service "InsightPulse" is used at the site "XXX.com", and a Cookie specifying InsightPulse has the Name attribute name "app1" and the domain attribute value "example.com".

[0090] The analysis item reference unit 82 is a functional unit which outputs, to the user terminal 90, a site URL (which may be a page URL of each page) and a cookie, web storage, or an external communication URL corresponding to the site URL. Note that the "output" referred to here may be in a form of output to the display 91 of the user terminal 90, in a form of transmission in a file format to the user terminal 90, or in any other form. The analysis item reference unit 82 includes an analysis item acquisition unit 821 and an analysis item output unit 822.

[0091] The analysis item acquisition unit (an example of a second analysis item acquisition unit) 821 is a functional unit which refers to the analysis item database 40 to acquire a site URL (which may be a page URL of each page) and a cookie, web storage, or an external communication URL corresponding to the site URL. For example, the analysis item acquisition unit 821 acquires, from a first row of the upper table of the analysis item database 40 illustrated in Fig. 5, analysis information in which a page URL is "XXX.com", a Name attribute name is "app1", and a domain attribute is "example.com".

[0092] The analysis item output unit 822 is a functional unit which outputs, to the user terminal 90, a cookie, web storage, or an external communication URL corresponding to the site URL (which may be a page URL of each page) acquired by the analysis item acquisition unit 821 .

[0093] Fig. 13 illustrates an example of a display screen on the display 91 in which the cookie, the web storage, or the external communication URL corresponding to the site URL (which may be a page URL of each page) acquired by the analysis item acquisition unit 821 is displayed. The screen includes a target site display unit 913 and an analysis item display unit 914. When the user specifies "XXX.com" as a reference target URL on the target site display unit 913, the reference apparatus 80 displays the site URL "XXX.com" on the target site display unit 913, and displays, on the analysis item display unit 914, the cookie, the web storage, or the external communication URL acquired as ones corresponding to the page URL "XXX.com" from the analysis item database 40 by

the analysis item acquisition unit 821 in association with one another. As a result, for example, it is displayed that a cookie with the Name attribute name "app1" and the domain attribute "example.com" has been issued on a top page "XXX.com" at the site "XXX.com".

[0094] The reference apparatus 80 according to the present embodiment includes: the using service reference unit 81 including the using service acquisition unit 811 which acquires a using service registered by the using service identifying apparatus 50 as a service used at an analysis target site, the analysis item acquisition unit 812 which acquires a cookie, web storage, or an external communication URL associated with the using service by referring to the service database 60, and the using service output unit 813 which outputs the using service, and the cookie, the web storage, or the external communication URL to the user terminal 90; and the analysis item reference unit 82 including the analysis item acquisition unit 812 which acquires an analysis item including a correspondence relationship between a site URL of the analysis target site and the cookie, the web storage, or the external communication URL corresponding to the site URL, and the analysis item output unit 822 which outputs, to the user terminal 90, the site URL included in the analysis item and the cookie, the web storage, or the external communication URL corresponding to the site URL. As a result, the service used at the target site can be output in association with the cookie for accessing the using service, the web storage, or the external communication URL. It is possible not only to list the using services, but also to provide the user with the correspondence relationship between the using services and the analysis items.

[0095] Fig. 14A illustrates a flow of a using service reference method executed by the reference apparatus 80.

[0096] In step S31, a target site that the user desires to refer to is specified. In the present example, the reference apparatus 80 specifies the target site from an input screen output to the display 91 of the user terminal 90. Fig. 14B illustrates a target input screen on which the user

specifies the target site. The analysis target input screen has an input form 915, and displays data of all rows of the site database 20 in a pull-down format so that the user can select the data.

[0097] In step S32, the reference apparatus 80 (the using service acquisition unit 811 of the using service reference unit 81) acquires, from the using service database 70 (see Fig. 9), all using services registered by the using service identifying method (see Fig. 13) as services used at the analysis target site specified by the user in step S31.

[0098] In step S33, the reference apparatus 80 (the analysis item acquisition unit (an example of the third analysis item acquisition unit) 812 of the using service reference unit 81) refers to the service database 60 (see Fig. 8B) and acquires a cookie, web storage, or an external communication URL by using a using service name as a key. For example, a cookie with the Name attribute name "app1" and the domain attribute value "example.com" corresponding to the service name "InsightPulse" is acquired. Details of the acquisition of the cookie or the like are as described above.

[0099] In step S34, the reference apparatus 80 (the using service output unit 813 of the using service reference unit 81) outputs, as a set, the using service together with the cookie, the web storage, or the external communication URL to a display of the user terminal 90 (see Fig. 12). For example, a cookie with the service name "InsightPulse" and the Name attribute name "app1" and the domain attribute value "example.com" corresponding to the service are output in one row (see Fig. 12).

[0100] The using service reference method according to the present embodiment is a method executed by a computer, and includes acquiring, as a service used at an analysis target site, a using service registered by the using service identifying method described above, acquiring a cookie, web storage, or an external communication URL associated with the using service by referring to a service database, and outputting the using service and the cookie, the web storage, or the external

communication URL. According to this, the service used at the target site can be output in association with the cookie, the web storage, or the external communication URL for accessing the using service. It is possible not only to list the using services in the site but also to provide the user with the using services and the analysis items in association with one another.

[0101] Fig. 15 illustrates a flow of an analysis information reference method executed by the reference apparatus 80.

[0102] In step S41, an analysis target site that the user desires to refer to is specified. In the present example, the reference apparatus 80 specifies the analysis target site from an input screen output to the display 91 of the user terminal 90. The analysis target input screen has an input form 915, and displays data of all rows of the site database 20 in a pull-down format (see Fig. 14B).

[0103] In step S42, the reference apparatus 80 (the analysis item acquisition unit 821 of the using service reference unit 81) acquires, from the analysis item database 40 (see Fig. 5), all the analysis items registered by the analysis method (see Fig. 6) as services used at the analysis target site. For example, the analysis item acquisition unit 821 acquires, from the first row of the upper table of the analysis item database 40 illustrated in Fig. 5, analysis information in which a site URL is "XXX.com", a Name attribute name is app1, and a domain attribute is example.com.

[0104] In step S43, the reference apparatus 80 (the analysis item output unit 822 of the using service reference unit 81) outputs, to the display of the user terminal 90, all analysis information in which the site URL acquired by the analysis item acquisition unit 821 is associated with a cookie, web storage, or an external communication URL in the site (see Fig. 13). For example, a cookie with the page URL "XXX.com", a Name attribute name "app1" and a domain attribute value "example.com" issued by this page is output in one row.

[0105] The analysis item reference method according to the present embodiment is executed by a computer, and includes acquiring, as analysis information of a page in an analysis target site, an

analysis item registered by the analysis method described above, and outputting all the acquired analysis information to the user terminal 90. According to this, the analysis item included in the target site can be output. A list of analysis items included in the target site can be provided to the user.

[0106] Note that the using service identifying system 1 according to the present embodiment includes the analysis apparatus 10, the using service identifying apparatus 50, the reference apparatus 80, and the service database 60 in which data in which each of a cookie, web storage, and an external communication URL is associated with a service name is registered. According to this, the analysis apparatus 10 extracts an analysis item, the using service identifying apparatus 50 matches the analysis item with the service database 60 to identify a using service of the site, and the reference apparatus 80 can provide the user with information regarding the using service.

[0107] In addition, the using service identifying system 1 according to the present embodiment includes: the analysis apparatus 10 which acquires a site URL of an analysis target site, analyzes one or more pages in a site published at the site URL, extracts cookies issued by the one or more pages, web storage set by the one or more pages, or external communication URLs of external communication generated by the one or more pages, and registers the site URL of the page, which issues the cookie, sets the web storage, or generates the external communication URL, and the cookie issued by the page, the web storage set by the page, or the external communication URL generated by the page in association with one another; the using service identifying apparatus 50 which matches an analysis item including a correspondence relationship between the site URL registered by the analysis apparatus 10 and the cookie, the web storage, or the external communication URL with the service database 60 to identify, as a using service used on the page at the site URL, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item and registers the site URL and the using service in

association with one another; and the service database 60 in which data in which each of a cookie, web storage, and an external communication URL is associated with a service is registered. According to this, it is possible to manage an external service used at the site and a cookie or the like issued to access the external service in association with one another.

[0108] In addition, the using service identifying system 1 according to the present embodiment further includes the analysis item database 40 in which a correspondence relationship between a site URL and a cookie, web storage, or an external communication URL is registered by the analysis apparatus, and an analysis item including the correspondence relationship is acquired by the analysis apparatus 10.

[0109] In addition, the using service identifying system 1 according to the present embodiment further includes the reference apparatus 80 which acquires, as the service used at the analysis target site, the using service registered by the using service identifying apparatus 50 and outputs the using service.

[0110] In addition, the using service identifying system 1 according to the present embodiment further includes the using service database 70 in which the site URL and the using service is registered in association with one another by the using service identifying apparatus 50, and the using service is acquired by the reference apparatus 80.

[0111] In addition, the using service identifying system 1 according to the present embodiment further includes the site database 20 in which a site URL of at least one analysis target site is registered, and the site URL is acquired by the analysis apparatus 10.

[0112] In addition, the using service identifying method according to the present embodiment includes: acquiring a site URL of an analysis target site, analyzing one or more pages in a site published at the site URL, extracting cookies issued by the one or more pages, web storage set by the one or more pages, or external communication URLs of external communication generated by

the one or more pages, and registering the site URL of the page, which issues the cookie, sets the web storage, or generates the external communication URL, and the cookie issued by the page, the web storage set by the page, or the external communication URL generated by the page in association with one another; and matching an analysis item including a correspondence relationship between the site URL registered by the registering and the cookie, the web storage, or the external communication URL with the service database 60 to identify, as a using service used on the page at the site URL, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item and registering the site URL and the using service in association with one another, and the service database 60 registers data in which each of a cookie, web storage, and an external communication URL is associated with a service. According to this, it is possible to manage an external service used at the site and a cookie or the like issued to access the external service in association with one another.

[0113] In addition, the program according to the present embodiment causes one or more computers to execute: acquiring a site URL of an analysis target site, analyzing one or more pages in a site published at the site URL, extracting cookies issued by the one or more pages, web storage set by the one or more pages, or external communication URLs of external communication generated by the one or more pages, and registering the site URL of the page, which issues the cookie, sets the web storage, or generates the external communication URL, and the cookie issued by the page, the web storage set by the page, or the external communication URL generated by the page in association with one another; and matching an analysis item including a correspondence relationship between the site URL registered by the registering and the cookie, the web storage, or the external communication URL with the service database 60 to identify, as a using service used on the page at the site URL, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item and registering the site URL and the using service

in association with one another, and the service database 60 registers data in which each of a cookie, web storage, and an external communication URL is associated with a service. According to this, it is possible to manage an external service used at the site and a cookie or the like issued to access the external service in association with one another.

[0114] (Second Embodiment)

Fig. 16 illustrates a schematic configuration of a using service identifying system (also simply referred to as a system) 1A of a second embodiment. The system 1A is a network system which analyzes a site while determining a security risk, to identify a using service used at the site, and provides information regarding the identified using service to, for example, the user terminal 90 or the like used by an administrator of a website. The system 1A includes a security risk determination system 2, the using service identifying apparatus 50, and the reference apparatus 80. The system 1A may further include the service database (service DB) 60 and the using service database (using service DB) 70. The security risk determination system 2 includes a proxy apparatus 100, the analysis apparatus 10, and a status reference apparatus 300. The security risk determination system 2 may further include a status database 200, the site database 20, and the analysis item database 40. These components are communicably connected to one another via the network 99. Note that apparatuses that are the same as or correspond to the apparatuses in the first embodiment (the analysis apparatus 10, the analysis target server 30, the using service identifying apparatus 50, the reference apparatus 80, the user terminal 90, the site database 20, the analysis item database 40, the service database 60, the using service database 70, the user terminal 90) are denoted by same reference numerals, and only differences from the first embodiment will be described in detail (contents that are the same as or similar to those of the first embodiment are incorporated by reference in the present embodiment to avoid repetition of description).

[0115] The proxy apparatus 100 is a computer apparatus which mediates between the analysis target server 30 and the analysis apparatus 10 (virtual browser), analyzes a response from the analysis target server 30 to determine a security risk, and transfers the response to the analysis apparatus 10 (virtual browser) when it is determined that there is no security risk. The proxy apparatus 100 includes a central processing unit (CPU) and a communication apparatus (both not illustrated). Note that the proxy apparatus 100 may be implemented by a cloud (a plurality of distributed servers, a plurality of subsystems, or the like). The CPU executes a security risk determination program to exhibit a function of analyzing a response from the analysis target server 30 to determine a security risk. A functional configuration of the proxy apparatus 100 will be described later. Note that the security risk determination program is activated, for example, by being stored in a ROM (not illustrated) and read by the CPU or being stored in a storage medium such as a CD-ROM and read by the CPU using a reading apparatus (not illustrated) to be developed in a RAM. The communication apparatus is an apparatus capable of performing wired or wireless communication with various apparatuses constituting the system 1A, various databases, and further the analysis target server 30 via the network 99, and can perform communication using a protocol such as TCP/IP as an example.

[0116] Fig. 17 illustrates a functional configuration of the proxy apparatus 100. The proxy apparatus 100 includes a request determination unit 105, a request transfer unit 101, a response analysis unit 102, a response return unit 103, a request blocking unit 106, and a response blocking unit 104.

[0117] The request determination unit 105 is a functional unit which receives a request transmitted from the virtual browser, determines a security risk of a request URL specified in the request, and determines that a request URL having a security risk is abnormal and a request URL without any security risk is normal. Here, the security risk may be determined, for example, by

inputting a request URL to an existing security determination external service such as GOOGLE (registered trademark) Safe Browsing and obtaining a security risk determination result.

[0118] The request transfer unit 101 is a functional unit which receives a request transmitted from the virtual browser and transfers the request to the analysis target server 30 which operates an analysis target site. Note that the analysis target site is a target site for which a service used at the analysis target site is identified by the system 1A according to the present embodiment. In addition, the virtual browser is a virtual apparatus which is activated by the analysis apparatus 10 in order to specify a site URL of an analysis target site, transmit a request to the analysis target server 30 which operates the analysis target site, and receive a response from the analysis target server 30. The analysis target server 30 returns a response to the proxy apparatus 100 in response to receiving a request.

[0119] The response analysis unit 102 is a functional unit which receives and analyzes a response from the analysis target server 30. This analysis includes detecting, from content in the response, a malicious script such as known malware, virus, or cross-site scripting by using a security signature database or the like, executing a code in the response (see the script 33 in Fig. 4) in a sandbox environment, confirming whether or not response header information is appropriate, or the like. When a malicious script is detected from the content in the response, when the code in the response indicates a malicious or invalid operation in the sandbox environment, when the response header information is not appropriate, or the like, it is determined that there is a security risk. In addition, when a status code included in status code information 35 of the response transmitted from the analysis target server 30 is not 200, the response analysis unit 102 determines that a page included in the response is a non-referable page.

[0120] The response return unit 103 is a functional unit which returns a response to the virtual browser. The response return unit 103 returns, to the virtual browser, a response determined by

the response analysis unit 102 to have no security risk. For example, when the response analysis unit 102 determines that the external domain "example.com" has no security risk, the response return unit 103 returns the data of Fig. 4 as it is as a response.

[0121] When the request determination unit 105 determines that the request URL is abnormal, the request blocking unit 106 blocks a request specifying the request URL. When it is determined that the request URL is abnormal, the request blocking unit 106 generates a new response (that is, an error response) indicating that the request URL has a security risk, and returns the new response to the virtual browser.

[0122] When the response analysis unit 102 determines that the response has a security risk, the response blocking unit 104 blocks communication with the analysis target site. The response blocking unit 104 further generates a new response (that is, an error response) indicating that the page for which the request has been transmitted by the request transfer unit 101 has a security risk, and returns the new response to the virtual browser.

[0123] Fig. 18 illustrates an example of data included in the error response. The HTTP/HTTPS response includes a response line (status line), a response header, and a response body. In the present example, the response line includes a version of a HTTP protocol and status code information 1001, and it is indicated that the version is HTTP/1.1, and the status code is 200. The response header includes generation date and time and connection information of the response, and indicates that the response is generated at 10:18:14 on June 9, 2024. The response body includes an HTML document, and the HTML document includes blocked external domain information 1002 "Blocked domain: example.com".

[0124] The proxy apparatus 100 according to the present embodiment includes: the request determination unit 105 which receives a request transmitted from a virtual browser and determines a security risk of a request URL specified in the request; the request transfer unit 101 which transfers the request for which it is determined by the request determination unit 105 that the request URL is normal, to the server 30 operating an analysis target site, the analysis target site

being a target site for which a service used at the analysis target site is identified; the response analysis unit 102 which receives and analyzes a response from the server 30; and the response return unit 103 which returns the response to the virtual browser, so that the analysis apparatus 10 can be prevented from receiving a response having a security risk.

[0125] Fig. 19 illustrates a flow of a security risk determination method S100 executed by the proxy apparatus 100.

[0126] In step S101, the proxy apparatus 100 (request determination unit 105) receives a request transmitted from the virtual browser activated by the analysis apparatus 10, and determines a security risk of a request URL specified by this request. Note that the virtual browser is a virtual apparatus which is activated by the analysis apparatus 10 in order to specify a site URL of an analysis target site, transmit a request to the analysis target server 30 which operates the analysis target site, and receive a response from the analysis target server 30. In addition, a method for determining the security risk is as described above.

[0127] In step S102, when the proxy apparatus 100 (request determination unit 105) determines that the request URL has a security risk, in other words, the request URL is abnormal, the process proceeds to step S103. When the proxy apparatus 100 (request determination unit 105) determines that the request URL has no security risk, in other words, the request URL is normal, the process proceeds to step S104.

[0128] In step S103, the proxy apparatus 100 (request blocking unit 106) blocks the request. The request blocking unit 106 further generates a new response (that is, an error response) indicating that the request URL has a security risk, and returns the new response to the virtual browser. For example, when the request determination unit 105 determines that the request URL "example.com" has a security risk, the request blocking unit 106 generates data of Fig. 18 and returns the data as the response.

[0129] In step S104, the proxy apparatus 100 (request transfer unit 101) transfers, to the analysis target server 30 which operates the analysis target site, a request for which the request determination unit 105 determines that the request URL is normal. Note that the analysis target site is a target site for which a service used at the analysis target site is identified by the system 1A

according to the present embodiment.

[0130] In step S105, the proxy apparatus 100 (response analysis unit 102) receives and analyzes the response from the analysis target server 30. Note that an analysis method is as described above.

[0131] In step S106, the proxy apparatus 100 (response analysis unit 102) determines presence or absence of a security risk in the response. If it is determined that there is a security risk, the process proceeds to step S107, and if it is determined that there is no security risk, the process proceeds to step S108.

[0132] In step S107, the proxy apparatus 100 (response blocking unit 104) blocks the response from the analysis target site. The response blocking unit 104 further generates a new response (that is, an error response) indicating that the page included in the response has a security risk, and returns the new response to the virtual browser.

[0133] In step S108, the proxy apparatus 100 (response return unit 103) returns the response to the virtual browser. For example, when the request determination unit 105 determines that the request URL is normal and the response analysis unit 102 does not determine that the response has a security risk, the response return unit 103 returns the response as it is.

[0134] The security risk determination method according to the present embodiment includes: receiving a request transmitted from a virtual browser and determining a security risk of a request URL specified in the request; requesting by transferring the request for which it is determined in the determining that the request URL is normal, to the analysis target server 30 operating an analysis target site, the analysis target site being a target site for which a service used at the analysis target site is identified; receiving and analyzing a response from the server; and returning the response to the virtual browser, so that the analysis apparatus 10 can be prevented from receiving a response having a security risk.

[0135] Fig. 20 illustrates a functional configuration of the analysis apparatus 10. Note that an apparatus configuration of the analysis apparatus 10 is as described above. The analysis apparatus 10 includes the analysis target acquisition unit 11, the analysis execution unit 12, the analysis item registration unit 13, and a status registration unit 14. Note that the analysis target

acquisition unit 11 and the analysis item registration unit 13 are similar to the analysis target acquisition unit 11 and the analysis item registration unit 13 according to the first embodiment described above.

[0136] The analysis execution unit 12 is a functional unit which analyzes a page having no security risk among one or more pages in a site published at a site URL for an analysis target site (for example, "XXX.com") acquired from the site database 20 by the analysis target acquisition unit 11, and extracts a cookie issued by each of the pages, web storage set by each of the pages, or an external communication URL of external communication generated by each of the pages. The analysis execution unit 12 includes the site access unit 121, a status extraction unit 125, the analysis item extraction unit 122, the script execution unit 123, and the link extraction unit 124.

[0137] The site access unit 121 is a functional unit which specifies a site URL of an analysis target site, transmits a request to the proxy apparatus 100, and receives a response from the analysis target server 30 which operates the analysis target site via the proxy apparatus 100 by using the virtual browser. The site access unit 121 further determines presence or absence of a link URL that has not yet been specified yet among link URLs extracted by the link extraction unit 124, and when there is a link URL that has not yet been specified and accessed, the site access unit specifies the link URL and transmits a request to the proxy apparatus 100, and receives a response from the analysis target server 30 which operates the analysis target site via the proxy apparatus 100. The site access unit 121 can recursively analyze all pages traced from a top page of the analysis target site by specifying the link URL, transmitting the request to the proxy apparatus 100, and receiving the response from the analysis target server 30 which operates the analysis target site via the proxy apparatus 100.

[0138] The status extraction unit 125 is a functional unit which extracts a security risk from the response received by the site access unit 121. The status extraction unit 125 may extract a status

code or the like of the response. When the response received by the site access unit 121 is an error response, the status extraction unit 125 determines that a page for which the request is transmitted by the site access unit 121 has a security risk. The status extraction unit 125 extracts the security risk and the status code from the response received by the site access unit 121, thereby acquiring a security risk, a broken link, and the like of the page. For example, when an error response (see Fig. 18) is transmitted from the proxy apparatus, it is acquired that the requested page has a security risk, from the information 1002 indicating that the external domain URL is determined to be malicious or invalid. Note that a method of extracting the status code or the like included in the response is not particularly limited, and for example, the status code or the like can be extracted by using an API such as JAVASCRIPT (registered trademark), Python, or JAVA (registered trademark).

[0139] The analysis item extraction unit 122 is a functional unit which when a status extracted by the status extraction unit 125 is normal (status code 200), extracts, from the response received by the site access unit 121, cookies issued by one or more pages or external communication URLs generated by the one or more pages and further acquires data stored in web storage set in the virtual browser. Details of the processing are similar to those of the analysis item extraction unit 122 in the first embodiment.

[0140] The script execution unit 123 is a functional unit which, when the status extracted by the status extraction unit 125 is normal (status code 200) and the response transmitted by the proxy apparatus 10 includes a script, executes the script by the virtual browser. The script is generally described in an HTML document, and the HTML document is included in a response body of an HTTP response (see the script tag 33 in Fig. 3). An execution result of the script can be confirmed by, for example, a log or the like output from an interface driver which operates the virtual browser. As a result, it is possible to confirm whether or not execution of the script is successful.

[0141] The link extraction unit 124 is a functional unit which, when the status extracted by the status extraction unit 125 is normal (status code 200), extracts a page URL of a linked page from the response transmitted by the proxy apparatus 100. Details of the processing are similar to those of the link extraction unit 124 in the first embodiment.

[0142] The status registration unit 14 registers, in the status database 200, page URLs (which may be a site URL) of one or more pages, and the status code and the security risk extracted by the status extraction unit 125. For example, when the site access unit 121 receives an error response (see Fig. 18) with respect to a request specifying the page URL "XXX.com", the status registration unit 14 registers, in the status database 200, the page URL "XXX.com", a status code "200", and a security risk "present (1)" in association with one another. When the site access unit 121 receives a response that the request specifying the page URL "XXX.com/XXX" is processed normally, the status registration unit 14 registers, in the status database 200, the page URL "XXX.com/XXX", the status code "200" indicating normality, and a security risk "absent (0)" in association with one another. When a response of a status code "404" is returned with respect to the request specifying the page URL "XXX.com/ZZZ", the site access unit 121 registers, in the status database 200, the page URL "XXX.com/ZZZ", the status code "404" indicating that this page has not been found, and the security risk "absent (0)" in association with one another. The status registration unit 14 which registers, in the status database 200, the page URL of the page and the status code and the security risk of the response to the request specifying the page URL in association with one another can record a security risk analysis result by the proxy apparatus 100 so as to be available to another apparatus.

[0143] The status database 200 is a storage apparatus in which status information in which a page URL of a page is associated with a status code of a response to a request specifying the page URL, and presence or absence of a security risk is registered (stored) by the analysis item

registration unit 13. A hard disk drive (HDD) or the like may be included as the storage apparatus.

[0144] Fig. 21 illustrates an example of the status information registered in the status database 200. The status information is registered in a table format including a page URL (URL) column, a status code column, and a security risk column. In the present example, for example, the page URL "XXX.com" is registered in association with the status code "200" indicating that a response is returned normally from this page and "1" indicating that there is a security risk, the page URL "XXX.com/XXX" is registered in association with the status code "200" indicating that a response is returned normally from this page and "0" indicating that there is no security risk, the page URL "XXX.com/YYY" is registered in association with the status code "200" of this page and "0" indicating that there is no security risk, and the page URL "XXX.com/ZZZ" is registered in association with the status code "404" indicating a broken link of this page and "0" indicating that there is no security risk. In addition to these, other information regarding the status may be registered in the status database 200. The other information is, for example, an external domain causing a security risk, a service name of a service specified by the external domain, or the like.

[0145] The analysis apparatus 10 according to the present embodiment includes the analysis target acquisition unit 11 which acquires a site URL of an analysis target site and the analysis execution unit 12 which analyzes one or more pages in a site published at the site URL, and the analysis execution unit 12 includes the site access unit 121 which specifies the site URL, transmits a request to the proxy apparatus 100, and receives a response from the proxy apparatus 100 by using the virtual browser, so that it is possible to receive only the response without any security risk and prevent infection with malware or the like. Then, similarly to the first embodiment, the using service identifying apparatus 50 can match an analysis item including a correspondence relationship between the site URL registered by the analysis apparatus 10 and the cookie, the web storage, or the external communication URL with the service database 60 to identify, as a using

service used on the page at the site URL, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item and registers the site URL and the using service in association with one another. In addition, the reference apparatus 80 can acquire, as the service used at the analysis target site, the using service registered by the using service identifying apparatus 50 and output the using service. Details are as described in the first embodiment.

[0146] Fig. 22 illustrates a flow of the analysis method executed by the analysis apparatus 10.

[0147] In step S11, the analysis apparatus 10 (analysis target acquisition unit 11) acquires site URLs of all analysis target sites from the site database 20. As described above, information (URL information) regarding the site URLs of the analysis target sites is registered in the site database 20. The URL information may be acquired by the operator of the system 1 inputting the URL information via the input device (not illustrated).

[0148] In step S12, the analysis apparatus 10 (analysis execution unit 12) analyzes a page in a site published at the site URL, and extracts all of a cookie issued by the page, web storage set by the page, or an external communication URL generated by the page. In addition, all link URLs are extracted. First, the site access unit 121 specifies a site URL (a page URL of a top page of the site) or a page URL of a linked page extracted by the link extraction unit 124 and transmits a request to the proxy apparatus 100 by using the virtual browser. The proxy apparatus 100 executes the security risk determination method S100 (see Fig. 19) in response to receiving the request, and the site access unit 121 receives a response from the analysis target server 30 operating the analysis target site via the proxy apparatus 100. Details of the security risk determination method S100 are as described above.

[0149] For example, when an error response (see Fig. 18) is transmitted from the proxy apparatus, "200" included in the status code 1001 is acquired as the status code, and it is acquired

that the requested page has a security risk, from the information 1002 indicating that the external domain URL is determined to be malicious or invalid. When a normal response (see Fig. 4) is transmitted via the proxy apparatus, "200" included in the status code information 35 is acquired as a status code, and since a response body does not include information indicating that the external domain URL is determined to be malicious or invalid, it is acquired that the requested page has no security risk. Details are as described above.

[0150] Next, the status extraction unit 125 extracts the security risk and the status code from the response received by the site access unit 121. When the page for which the request is transmitted by the site access unit 121 has no security risk, the analysis item extraction unit 122 further extracts, from the response received by the site access unit 121, cookies issued by one or more pages or external communication URLs generated by one or more pages, and further acquires data stored in web storage set in the virtual browser. Thereafter, when a script is included in the response transmitted by the proxy apparatus 100, the script execution unit 123 executes the script by the virtual browser. Then, the link extraction unit 124 extracts the page URL of the linked page from the response transmitted by the proxy apparatus 100.

[0151] In step S13, the analysis apparatus 10 (site access unit 121) determines presence or absence of a link URL that has not yet been specified, among the link URLs extracted by the link extraction unit 124. If there is a link URL that has not yet been specified, the process returns to step S12, and if not, the process proceeds to step S14.

[0152] In step S14, the analysis apparatus 10 (site access unit 121) determines presence or absence of a site URL that has not yet been specified, among the site URLs extracted by the analysis target acquisition unit 11. If there is a site URL that has not yet been specified, for example, if analysis of "XXX.com" (see Fig. 3) as the analysis target site has been completed and analysis of "YYY.com" and "ZZZ.com" (see Fig. 3) as the analysis target sites has not yet been completed, the

process returns to step S12, and if there is not the site URL (if the analysis of all the analysis target sites has been completed), the process proceeds to step S15.

[0153] In step S15, the analysis apparatus 10 (analysis item registration unit 13) associates the page URL of the page issuing the cookie, the web storage, or the external communication URL with the cookie, the web storage, or the external communication URL, and registers them in the analysis item database 40 (see Fig. 5) so as to be acquirable by the using service identifying apparatus 50 (acquirable by the using service identifying method to be described later). Details of the accumulation of the analysis information are as described above.

[0154] In step S16, the analysis apparatus 10 (status registration unit 14) registers, in the status database 200, each of page URLs of one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs in association with one another. As a result, for example, when the site access unit 121 receives an error response (see Fig. 18) with respect to a request specifying the page URL "XXX.com", the status registration unit 14 registers, in the status database 200, the page URL "XXX.com", the status code "403", and the security risk "present (1)" in association with one another. When the site access unit 121 receives a response that the request specifying the page URL "XXX.com/XXX" is processed normally, the status registration unit 14 registers, in the status database 200, the page URL "XXX.com/XXX", the status code "200" indicating normality, and the security risk "absent (0)" in association with one another. When a response of the status code "404" is returned with respect to the request specifying the page URL "XXX.com/ZZZ", the site access unit 121 registers, in the status database 200, the page URL "XXX.com/ZZZ", the status code "404" indicating that this page has not been found, and the security risk "absent (0)" in association with one another. In this manner, the status information is accumulated in the status database 200. Details of the accumulation of the status information are as described above.

[0155] The analysis method according to the present embodiment includes acquiring a site URL of an analysis target site and analyzing one or more pages in a site published at the site URL, and the analyzing includes specifying the site URL, executing the security risk determination method, and receiving a response by using the virtual browser, and extracting a security risk from the response, so that the analysis apparatus 10 can receive only the response without any security risk and prevent infection with malware or the like. Then, similarly to the first embodiment, by the using service identifying method, it is possible to match an analysis item including a correspondence relationship between the site URL registered by the analysis method and the cookie, the web storage, or the external communication URL with the service database 60 to identify, as a using service used on the page at the site URL, a service corresponding to the cookie, the web storage, or the external communication URL included in the analysis item and register the site URL and the using service in association with one another. In addition, by the reference method, the using service registered by the using service identifying method can be acquired as the service used at the analysis target site, and the using service can be output. Details are as described in the first embodiment.

[0156] The reference apparatus (status reference apparatus) 300 is a computer apparatus which outputs a page URL of a page in a target site in association with a status such as a security risk of the page. The reference apparatus 300 includes a CPU and a communication apparatus (both not illustrated). Note that the reference apparatus 300 may be implemented by a cloud (a plurality of distributed servers, a plurality of subsystems, or the like).

[0157] Fig. 23 illustrates a functional configuration of the reference apparatus 300. The reference apparatus 300 includes a status acquisition unit 301 and a status output unit 302.

[0158] The status acquisition unit 301 is a functional unit which acquires, from the status database 200 registered by the analysis apparatus 10, page URLs of one or more pages and a status code of a response to a request specifying each of the page URLs and a security risk of each page. For example, the status acquisition unit 301 acquires, from the status database 200 (see Fig. 21), data associating the page URL "XXX.com/" with the status code "200" and the security risk "present (1)".

[0159] The status output unit 302 is a functional unit which outputs, to the user terminal 90, page URLs of one or more pages and a status code of a response to a request specifying each of the page URLs and a security risk of each page. For example, the status output unit 302 outputs data in which the page URL "XXX.com" acquired by the status acquisition unit 301 is associated with the status code "200" and the security risk "present (1)". Note that the "output" referred to here may be in a form of output to the display 91 of the user terminal 90, in a form of transmission in a file format to the user terminal 90, or in any other form.

[0160] Fig. 24 illustrates an example of the display screen of the display 91 on which the page URL of each page and the status code and the security risk of the response to the request specifying the page URL, which are acquired by the status acquisition unit 301, are displayed. The screen includes a target site display unit 916 and a status information display unit 917. When the user specifies "XXX.com" as a reference target URL on the target site display unit 913, the reference apparatus 300 displays the site URL "XXX.com" on the target site display unit 916, and displays, on the status information display unit 917, the status code and the security risk acquired as ones corresponding to the page URL "XXX.com/" from the status database 200 by the status acquisition unit 301 in association with one another . As a result, for example, the status code of the response to the request specifying the site "XXX.com" is "200", and it is displayed on this page that there is a security risk.

[0161] The reference apparatus 300 according to the present embodiment includes the status acquisition unit 301 which acquires page URLs of one or more pages, a status code of a response to a request specifying each of the page URLs, and a security risk of the pages, from the status database 200 registered by the analysis apparatus 10; and the status output unit which outputs, to the user terminal 90, the page URLs of the one or more pages, the status code of the response to the request specifying each of the page URLs, and the security risk of each page, so that it is possible to notify the administrator or the like of the analysis target site of the security risk of the page in the site, the broken link of the page, and the like.

[0162] Fig. 25 illustrates a flow of a reference method executed by the reference apparatus 300.

[0163] In step S301, an analysis target site that the user desires to refer to is specified. In the present example, the reference apparatus 300 specifies the analysis target site from an input screen output to the display 91 of the user terminal 90. The analysis target input screen has an input form on the target site display unit 916, and displays data of all rows of the site database 20 in a pull-down format (see Fig. 24). Here, it is assumed that the user selects the page URL "XXX.com" of the analysis target site.

[0164] In step S302, the reference apparatus 300 (status acquisition unit 301) acquires, from the status database 200 registered by the analysis apparatus 10, page URLs of one or more pages and a status code of a response to a request specifying each of the page URLs and security risk of each page. For example, the status acquisition unit 301 acquires, from the status database 200 (Fig. 21), data associating the page URL "XXX.com/" with the status code "200" and the security risk "present (1)". Details of the processing are as described above.

[0165] In step S303, the reference apparatus 300 (status output unit 302) outputs the page URLs of one or more pages and the status code of the response to the request specifying each of the page URLs and the security risk of each page (referred to as security risk information) to the display 91 of the user terminal 90 (see the status information display unit 917 in Fig. 24). For example, the status output unit 302 outputs data associating the page URL "XXX.com/" with the status code "200" and the security risk "present (1)". Note that the security risk information is not limited to being output to the display 91 of the user terminal 90, and may be transmitted to the user terminal 90 in a file format or in any other form. Details of the processing are as described above.

[0166] The reference method according to the present embodiment includes acquiring page URLs of one or more pages, a status code of a response to a request specifying the page URLs, and a security risk of the pages, from the status database 200 registered by the analysis apparatus 10; and outputting, to the user terminal 90, the page URLs of the one or more pages, the status code of the response to the request specifying each of the page URLs, and the security risk of each page, so that it is possible to notify the administrator or the like of the analysis target site of the security risk of the page in the site, the broken link of the page, and the like.

[0167] Various embodiments of the present invention may be described with reference to flowcharts and block diagrams, where the blocks may represent (1) a stage of a process in which an operation is performed or (2) a section of an apparatus that is responsible for performing the operation. Certain steps and sections may be implemented by a dedicated circuit, a programmable circuit provided with computer-readable instructions stored on a computer-readable medium, and/or a processor provided with the computer-readable instructions stored on the computer-readable medium. The dedicated circuit may include a digital and/or analog hardware circuit, and may include an integrated circuit (IC) and/or a discrete circuit. The programmable circuit may include a reconfigurable hardware circuit including memory elements such as logic AND, logic OR, logic XOR, logic NAND, logic NOR, and other logic operations, flip-flops,

registers, field programmable gate arrays (FPGA), programmable logic arrays (PLA), and the like.

[0168] The computer-readable medium may include any tangible device that can store instructions to be executed by a suitable device, and as a result, the computer-readable medium having instructions stored thereon includes a product including instructions that can be executed in order to create means for executing operations specified in the flowcharts or block diagrams. Examples of the computer-readable medium may include an electronic storage medium, a magnetic storage medium, an optical storage medium, an electromagnetic storage medium, a semiconductor storage medium, and the like. More specific examples of the computer-readable medium may include a FLOPPY (registered trademark) disk, a diskette, a hard disk, a random-access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or flash memory), an electrically erasable programmable read-only memory (EEPROM), a static random access memory (SRAM), a compact disc read-only memory (CD-ROM), a digital versatile disk (DVD), a BLU-RAY (registered trademark) disk, a memory stick, an integrated circuit card, and the like.

[0169] The computer-readable instruction may include: an assembler instruction, an instruction-set-architecture (ISA) instruction; a machine instruction; a machine dependent instruction; a microcode; a firmware instruction; state-setting data; or either a source code or an object code described in any combination of one or more programming languages, including an object oriented programming language such as SMALLTALK (registered trademark), JAVA (registered trademark), C++, or the like, and a conventional procedural programming language such as a "C" programming language or a similar programming language.

[0170] The computer-readable instructions may be provided for a processor or programmable circuit of a programmable data processing apparatus such as a computer locally or via a wide area network (WAN) such as a local area network (LAN), the Internet, or the like, and execute the

computer-readable instructions in order to create means for executing the operations designated in flowcharts or block diagrams. Here, the computer may be a personal computer (PC), a tablet computer, a smartphone, a work station, a server computer, a general purpose computer, a special purpose computer, or the like, and may be a computer system in which a plurality of computers is connected. Such a computer system in which a plurality of computers is connected is also referred to as a distributed computing system, and is a computer in a broad sense. In the distributed computing system, each of the plurality of computers executes a part of a program and passes data during execution of the program between the computers as necessary, so that the plurality of computers collectively executes the program.

[0171] Examples of the processor include a computer processor, a central processing unit (CPU), a processing unit, a microprocessor, a digital signal processor, a controller, a microcontroller, and the like. The computer may include one processor or a plurality of processors. In a multiprocessor system including a plurality of processors, each of the processors executes a part of a program, and passes data during execution of the program between the processors as necessary, so that the plurality of processors collectively execute the program. For example, in execution of multitasking, each of the plurality of processors may finely execute a part of each task by performing task switching for each time slice. In this case, which part of one program is executed by each processor dynamically changes. Which part of the program is executed by each of the plurality of processors may be statically determined by programming that takes multiprocessors into account.

[0172] Fig. 26 illustrates an example of a computer 1200 in which a plurality of aspects of the present invention may be embodied entirely or partially. A program installed in the computer 1200 can cause the computer 1200 to function as an operation associated with the apparatuses according to the embodiments of the present invention or as one or more sections of the

apparatuses, or can cause the operation or the one or more sections to be executed, and/or can cause the computer 1200 to execute a process according to the embodiments of the present invention or a stage of the process. Such programs may be executed by a CPU 1212 to cause the computer 1200 to perform specific operations associated with some or all of the blocks in the flowcharts and block diagrams described in the present specification.

[0173] The computer 1200 according to the present embodiment includes the CPU 1212, a RAM 1214, a graphics controller 1216, and a display device 1218, which are mutually connected by a host controller 1210. The computer 1200 also includes input/output units such as a communication interface 1222, a storage apparatus 1224 such as a hard disk drive, a DVD-ROM drive 1226, and an IC card drive, which are connected to the host controller 1210 via an input/output controller 1220. The computer also includes legacy input/output units such as an ROM 1230 and a keyboard 1242, which are connected to the input/output controller 1220 via an input/output chip 1240.

[0174] The CPU 1212 operates according to programs stored in the ROM 1230 and the RAM 1214, thereby controlling each unit. The graphics controller 1216 acquires image data generated by the CPU 1212 in a frame buffer or the like provided in the RAM 1214 or in itself, such that the image data is displayed on the display device 1218.

[0175] The communication interface 1222 communicates with other electronic devices via a network. The storage apparatus 1224 stores programs and data used by the CPU 1212 in the computer 1200. The DVD-ROM drive 1226 reads programs or data from a DVD-ROM 1227 and provides the programs or data to the storage apparatus 1224 via the RAM 1214. The IC card drive reads programs and data from an IC card, and/or writes the programs and data to the IC card.

[0176] The ROM 1230 stores therein boot programs and the like executed by the computer 1200 at the time of activation, and/or programs that depend on the hardware of the computer 1200.

The input/output chip 1240 may also connect various input/output units to the input/output controller 1220 via a parallel port, a serial port, a keyboard port, a mouse port, or the like.

[0177] Programs are provided by a computer-readable medium such as the DVD-ROM 1227 or the IC card. The programs are read from the computer-readable medium, are installed in the storage apparatus 1224, the RAM 1214, or the ROM 1230, which are also an example of the computer-readable medium, and are executed by the CPU 1212. Information processing described in these programs is read by the computer 1200, and provides cooperation between the programs and the various types of hardware resources described above. An apparatus or method may be constructed by realizing the operation or processing of information according to the use of the computer 1200.

[0178] For example, when communication is performed between the computer 1200 and an external device, the CPU 1212 may execute a communication program loaded in the RAM 1214 and instruct the communication interface 1222 to perform communication processing based on processing written in the communication program. Under the control of the CPU 1212, the communication interface 1222 reads transmission data stored in a transmission buffer processing region provided in a recording medium such as the RAM 1214, the storage apparatus 1224, the DVD-ROM 1227, or the IC card, transmits the read transmission data to the network, or writes reception data received from the network in a reception buffer processing region or the like provided on the recording medium.

[0179] In addition, the CPU 1212 may cause the RAM 1214 to read all or a necessary part of a file or database stored in an external recording medium such as the storage apparatus 1224, the DVD-ROM drive 1226 (DVD-ROM 1227), the IC card, or the like, and may execute various types of processing on data on the RAM 1214. Then, the CPU 1212 writes the processed data back to the external recording medium.

[0180] Various types of information such as various types of programs, data, tables, and databases may be stored in a recording medium and subjected to information processing. The CPU 1212 may execute, on the data read from the RAM 1214, various types of processing including various types of operations, information processing, conditional judgement, conditional branching, unconditional branching, information retrieval/replacement, or the like described throughout the present disclosure and specified by instruction sequences of the programs, and writes the results back to the RAM 1214. In addition, the CPU 1212 may search for information in a file, a database, or the like, in the recording medium. For example, when a plurality of entries, each having an attribute value of a first attribute associated with an attribute value of a second attribute, is stored in the recording medium, the CPU 1212 may retrieve, out of the plurality of entries, an entry with the attribute value of the first attribute specified that meets a condition, read the attribute value of the second attribute stored in said entry, and thereby acquiring the attribute value of the second attribute associated with the first attribute meeting a predetermined condition.

[0181] The programs or software modules described above may be stored in a computer-readable medium on or near the computer 1200. In addition, a recording medium such as a hard disk or a RAM provided in a server system connected to a dedicated communication network or the Internet can be used as a computer-readable medium, thereby providing a program to the computer 1200 via the network.

[0182] While the present invention has been described by way of the embodiments, the technical scope of the present invention is not limited to the above-described embodiments. It is apparent to persons skilled in the art that various alterations or improvements can be made to the above-described embodiments. It is also apparent from the scope of the claims that the embodiments added with such alterations or improvements can be included in the technical scope of the present invention.

[0183] It should be noted that each process of the operations, procedures, steps, stages, and the like performed by the device, system, program, and method shown in the claims, specification, or drawings can be executed in any order as long as the order is not indicated by “prior to”, “before”, or the like and as long as the output from a previous process is not used in a later process. Even if the operation flow is described using phrases such as "first" or "next" for the sake of convenience in the claims, specification, or drawings, it does not necessarily mean that the process must be performed in this order.

[0184] According to the present specification, configurations described in the following items are also disclosed.

(Item 1)

A proxy apparatus including:

a request determination unit which receives a request transmitted from a virtual browser and determines a security risk of a request URL specified in the request;

a request transfer unit which transfers, to a server operating an analysis target site, the request for which the request determination unit determines that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified;

a response analysis unit which receives and analyzes a response from the server; and

a response return unit which returns a response to the virtual browser.

(Item 2)

The proxy apparatus according to item 1, wherein

the response analysis unit determines a security risk from the response transmitted from the server, and

the response return unit returns, to the virtual browser, the response determined by the response analysis unit to have no security risk.

(Item 3)

The proxy apparatus according to item 1, further including a request blocking unit which blocks the request when the request determination unit determines that the request URL is

abnormal.

(Item 4)

The proxy apparatus according to item 3, wherein when it is determined that the request URL is abnormal, the request blocking unit generates a new response indicating that the request URL has a security risk and returns the new response to the virtual browser.

(Item 5)

The proxy apparatus according to item 2, further including a response blocking unit which blocks return of the response when the response analysis unit determines that the response has a security risk.

(Item 6)

The proxy apparatus according to item 5, wherein when the response analysis unit determines that the response has a security risk, the response blocking unit further generates a new response indicating that a page included in the response has a security risk and returns the new response to the virtual browser.

(Item 7)

The proxy apparatus according to item 1, wherein when a status code of the response transmitted from the server is not 200, the response analysis unit determines a page included in the response as a non-referable page.

(Item 8)

The proxy apparatus according to item 1, wherein the virtual browser is activated by an analysis apparatus in order to specify a site URL of the analysis target site, transmit the request to a server operating the analysis target site, and receive the response from the server.

(Item 9)

An analysis apparatus including:

an analysis target acquisition unit which acquires a site URL of an analysis target site; and

an analysis execution unit which analyzes one or more pages in a site published at the site URL, wherein

the analysis execution unit includes

a site access unit which specifies the site URL, transmits the request to the proxy apparatus according to item 1, and receives the response from the proxy apparatus by using the virtual browser, and

a status extraction unit which extracts a security risk from the response.

(Item 10)

The analysis apparatus according to item 9, wherein

the analysis execution unit further includes a link extraction unit which extracts a page URL of a linked page from the response, and

the site access unit further specifies a page URL of the linked page, transmits a request to the proxy apparatus, and receives a response from the proxy apparatus.

(Item 11)

The analysis apparatus according to item 9, further including a status registration unit which registers, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 12)

A reference apparatus including:

a status acquisition unit which acquires page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by the analysis apparatus according to item 11; and

a status output unit which outputs, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

(Item 13)

A security risk determination method executed by a computer, including:

receiving a request transmitted from a virtual browser and determining a security risk of a request URL specified in the request;

requesting by transferring, to a server operating an analysis target site, the request for which it is determined in the determining that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified;

receiving and analyzing a response from the server; and

returning a response to the virtual browser.

(Item 14)

An analysis method executed by a computer, including:

acquiring a site URL of an analysis target site; and

analyzing one or more pages in a site published at the site URL, wherein

the analyzing includes

specifying the site URL, transmitting the request to a proxy apparatus including a computer which executes the security risk determination method according to item 13, and receiving the response from the proxy apparatus by using the virtual browser, and

extracting a security risk from the response.

(Item 15)

The analysis method according to item 14, further including registering, in a status database, each of page URLs of the one or more pages and a status code and a security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 16)

A reference method executed by a computer, including:

acquiring page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by the analysis method according to item 15; and

outputting, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

(Item 17)

A security risk determination program which causes a computer to execute:

receiving a request transmitted from a virtual browser and determining a security risk of a request URL specified in the request;

transferring, to a server operating an analysis target site, the request for which it is determined in the determining that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified;

receiving and analyzing a response from the server; and

returning a response to the virtual browser.

(Item 18)

An analysis program which causes a computer to execute:

acquiring a site URL of an analysis target site; and

analyzing one or more pages in a site published at the site URL, wherein

the analyzing causes the computer to execute

specifying the site URL, transmitting the request to a proxy apparatus including a computer which executes the security risk determination program according to item 17, and receiving the response from the proxy apparatus by using the virtual browser, and

extracting a security risk from the response.

(Item 19)

The analysis program according to item 18, which causes the computer to further execute

registering, in a status database, each of page URLs of the one or more pages and a status code and a security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 20)

An analysis program which causes a computer to execute:

acquiring page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by executing the analysis program according to item 19; and

outputting, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

(Item 21)

A security risk determination system including:

the proxy apparatus according to item 1;

the analysis apparatus according to item 9;

the reference apparatus according to item 12; and

a status database in which data in which page URLs of one or more pages are associated with a status code and a security risk of a response to a request specifying each of the page URLs is registered.

(Item 22)

A security risk determination system including:

an analysis apparatus which identifies a service used at an analysis target site; and

a proxy apparatus which mediates a server operating the analysis target site and a virtual browser activated by the analysis apparatus, the analysis target site being a target site for which a service used at the analysis target site is identified, the security risk determination system executing

by the analysis apparatus, acquiring a site URL of the analysis target site, and specifying the site URL and transmitting a request to the proxy apparatus by using a virtual browser,

receiving, by the proxy apparatus, the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request,

transferring, by the proxy apparatus, the request for which it is determined that the request URL is normal, to a server operating an analysis target site,

receiving and analyzing, by the proxy apparatus, a response from the server,

returning, by the proxy apparatus, a response to the virtual browser, and

by the analysis apparatus, receiving the response from the proxy apparatus, extracting a security risk from the response, analyzing one or more pages in a site published at the site URL based on the response, and registering, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 23)

The security risk determination system according to item 22, wherein when it is determined that the request URL is abnormal, the proxy apparatus generates a new response indicating that the request URL has a security risk and returns the new response to the virtual browser.

(Item 24)

The security risk determination system according to item 22, wherein the proxy apparatus further executes blocking communication with the analysis target site when it is determined that the response has a security risk.

(Item 25)

The security risk determination system according to item 24, wherein in the blocking, when it is determined that the response has a security risk, the proxy apparatus further generates a new response indicating that a page included in the response has a security risk and returns the new response to the virtual browser.

(Item 26)

The security risk determination system according to item 22, further executing outputting, by a reference apparatus, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs to a user terminal from the status database registered in the registering.

(Item 27)

A security risk determination method executed by one or more apparatuses, the security risk determination method including:

by an analysis apparatus, acquiring a site URL of an analysis target site, and specifying the site URL and transmitting a request to a proxy apparatus by using a virtual browser, the analysis target site being a target site for which a service used at the analysis target site is identified;

receiving, by the proxy apparatus, the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request;

transferring, by the proxy apparatus, the request for which it is determined that the request URL is normal, to a server operating an analysis target site;

receiving and analyzing, by the proxy apparatus, a response from the server;

returning, by the proxy apparatus, a response to the virtual browser; and

by the analysis apparatus, receiving the response from the proxy apparatus, extracting a security risk from the response, analyzing one or more pages in a site published at the site URL based on the response, and registering, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

(Item 28)

A security risk determination program which causes one or more computers to execute:

acquiring a site URL of an analysis target site, and specifying the site URL and transmitting a request to a proxy apparatus by using a virtual browser, the analysis target site being a target site for which a service used at the analysis target site is identified;

receiving, by the proxy apparatus, the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request;

transferring, by the proxy apparatus, the request for which it is determined that the request URL is normal, to a server operating an analysis target site;

receiving and analyzing, by the proxy apparatus, a response from the server;

returning, by the proxy apparatus, a response to the virtual browser; and

receiving the response from the proxy apparatus, extracting a security risk from the response, analyzing one or more pages in a site published at the site URL based on the response, and registering, in a status database, each of page URLs of the one or more pages and a status code and the security risk of the response to the request specifying each of the page URLs in association with one another.

EXPLANATION OF REFERENCES

[0185] 1, 1A: system (using service identifying system); 2: security risk determination system; 10: analysis apparatus; 11: analysis target acquisition unit; 12: analysis execution unit; 13: analysis item registration unit; 14: status registration unit; 20: site database; 30: analysis target server; 31: cookie setting; 32: external communication URL setting; 33: script tag; 34: link tag; 35: status code information; 40: analysis item database; 50: using service identifying apparatus; 51: analysis item acquisition unit; 52: using service identifying unit; 53: using service registration unit; 60: service database; 70: using service database; 80: reference apparatus; 81: using service reference unit; 82: analysis item reference unit; 90: user terminal; 91: display; 99: network; 100: proxy apparatus; 101: request transfer unit; 102: response analysis unit; 103: response return unit; 104: response blocking unit; 105: request determination unit; 106: request blocking unit; 121: site access unit; 122: analysis item extraction unit; 123: script execution unit; 124: link extraction unit; 125: status extraction unit; 200: status database; 300: status reference apparatus; 301: status acquisition unit; 302: status output unit; 521: external communication URL matching unit; 522: cookie matching unit; 523: web storage matching unit; 811: using service acquisition unit; 812: analysis item acquisition unit (third analysis item acquisition unit); 813: using service output unit; 821: analysis item acquisition unit (second analysis item acquisition unit); 822: analysis item output unit; 911: target site display unit; 912: using service display unit; 913: target site display unit; 914: analysis item display unit; 915: input form; 916: target site display unit; 917: status information display unit; 1001: status code information; 1002: blocked external domain information; 1200: computer; 1210: host controller; 1212: CPU; 1214: RAM; 1216: graphics controller; 1218: display device; 1220: input/output

controller; 1222: communication interface; 1224: storage apparatus; 1226: DVD-ROM drive; 1240: input/output chip; 1242: keyboard; and S100: security risk determination method.

WHAT IS CLAIMED IS:

1. An analysis apparatus comprising:

an analysis target acquisition unit which acquires a site URL of an analysis target site; and

an analysis execution unit which analyzes one or more pages in a site published at the site URL, wherein

the analysis execution unit includes

a site access unit which specifies the site URL, transmits a request to a proxy apparatus, and receives a response from the proxy apparatus by using a virtual browser, and

a status extraction unit which extracts a security risk from the response, and

the proxy apparatus includes

a request determination unit which receives the request transmitted from the virtual browser and determines a security risk of a request URL specified in the request,

a request transfer unit which transfers, to a server operating an analysis target site, the request for which the request determination unit determines that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified,

a response analysis unit which receives and analyzes a response from the server, and

a response return unit which returns a response to the virtual browser.

2. The analysis apparatus according to claim 1, wherein

the response analysis unit determines a security risk from the response transmitted from the server, and

the response return unit returns, to the virtual browser, the response determined by the response analysis unit to have no security risk.

3. The analysis apparatus according to claim 1, further comprising a request blocking unit which blocks the request when the request determination unit determines that the request URL is abnormal.

4. The analysis apparatus according to claim 3, wherein when it is determined that the request

URL is abnormal, the request blocking unit generates a new response indicating that the request URL has a security risk and returns the new response to the virtual browser.

5. The analysis apparatus according to claim 2, further comprising a response blocking unit which blocks return of the response when the response analysis unit determines that the response has a security risk.

6. The analysis apparatus according to claim 5, wherein when the response analysis unit determines that the response has a security risk, the response blocking unit further generates a new response indicating that a page included in the response has a security risk and returns the new response to the virtual browser.

7. The analysis apparatus according to claim 1, wherein when a status code of the response transmitted from the server is not 200, the response analysis unit determines a page included in the response as a non-referable page.

8. The analysis apparatus according to claim 1, wherein the virtual browser is activated by an analysis apparatus in order to specify a site URL of the analysis target site, transmit the request to a server operating the analysis target site, and receive the response from the server.

9. The analysis apparatus according to claim 1, wherein

the analysis execution unit further includes a link extraction unit which extracts a page URL of a linked page from the response, and

the site access unit further specifies a page URL of the linked page, transmits a request to the proxy apparatus, and receives a response from the proxy apparatus.

10. The analysis apparatus according to claim 1, further comprising a status registration unit which registers, in a status database, each of page URLs of the one or more pages and a status code

and the security risk of the response to the request specifying each of the page URLs in association with one another.

11. A reference apparatus comprising:

a status acquisition unit which acquires page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by the analysis apparatus according to claim 10; and

a status output unit which outputs, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

12. An analysis method executed by a computer, comprising:

acquiring a site URL of an analysis target site; and

analyzing one or more pages in a site published at the site URL, wherein

the analyzing includes

specifying the site URL, transmitting a request to a proxy apparatus including a computer which executes a security risk determination method, and receiving a response from the proxy apparatus by using the virtual browser, and

extracting a security risk from the response, and

the security risk determination method includes

receiving the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request,

requesting by transferring, to a server operating an analysis target site, the request for which it is determined in the determining that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified,

receiving and analyzing a response from the server, and

returning a response to the virtual browser.

13. The analysis method according to claim 12, further comprising registering, in a status database, each of page URLs of the one or more pages and a status code and a security risk of the response to the request specifying each of the page URLs in association with one another.

14. A reference method executed by a computer, comprising:

acquiring page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by the analysis method according to claim 13; and

outputting, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

15. An analysis program which causes a computer to execute:

acquiring a site URL of an analysis target site; and

analyzing one or more pages in a site published at the site URL, wherein

the analyzing causes the computer to execute

specifying the site URL, transmitting a request to a proxy apparatus including a computer which executes a security risk determination program, and receiving a response from the proxy apparatus by using the virtual browser, and

extracting a security risk from the response, and

the security risk determination program causes the computer to execute

receiving the request transmitted from the virtual browser and determining a security risk of a request URL specified in the request,

transferring, to a server operating an analysis target site, the request for which it is determined in the determining that the request URL is normal, the analysis target site being a target site for which a service used at the analysis target site is identified,

receiving and analyzing a response from the server, and

returning a response to the virtual browser.

16. The analysis program according to claim 15, which causes the computer to further execute registering, in a status database, each of page URLs of the one or more pages and a status code and a security risk of the response to the request specifying each of the page URLs in association with one another.

17. A reference program which causes a computer to execute:

acquiring page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs, from the status database registered by executing the analysis program according to claim 16; and

outputting, to a user terminal, the page URLs of the one or more pages and the status code and the security risk of the response to the request specifying each of the page URLs.

18. A security risk determination system comprising:

the analysis apparatus according to claim 1;

the reference apparatus according to claim 11; and

a status database in which data in which page URLs of one or more pages are associated with a status code and a security risk of a response to a request specifying each of the page URLs is registered.