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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number
WO 01/25882 A1

(51) International Patent Classification[7]: G06F 1/00, 12/14

(21) International Application Number: PCT/GB00/03620

(22) International Filing Date:
21 September 2000 (21.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9923340.5 4 October 1999 (04.10.1999) GB

(71) Applicant *(for all designated States except US)*: **THE SECRETARY OF STATE FOR DEFENCE** [GB/GB]; Defence Evaluation Research Agency, A4 Building, Ively Road, Farnborough, Hampshire GU14 0LX (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SIMPSON, Gary, Steven** [GB/GB]; DERA Malvern, St Andrews Road, Malvern, Worcestershire WR14 3PS (GB). **DEAN, Timothy, Barry** [GB/GB]; DERA Malvern, St Andrews Road, Malvern, Worcestershire WR14 3PS (GB).

(74) Agent: **BOWDERY, A., O.**; D/IPD, DERA Formalities, A4 Building, Ively Road, Farnborough, Hampshire GU14 0LX (GB).

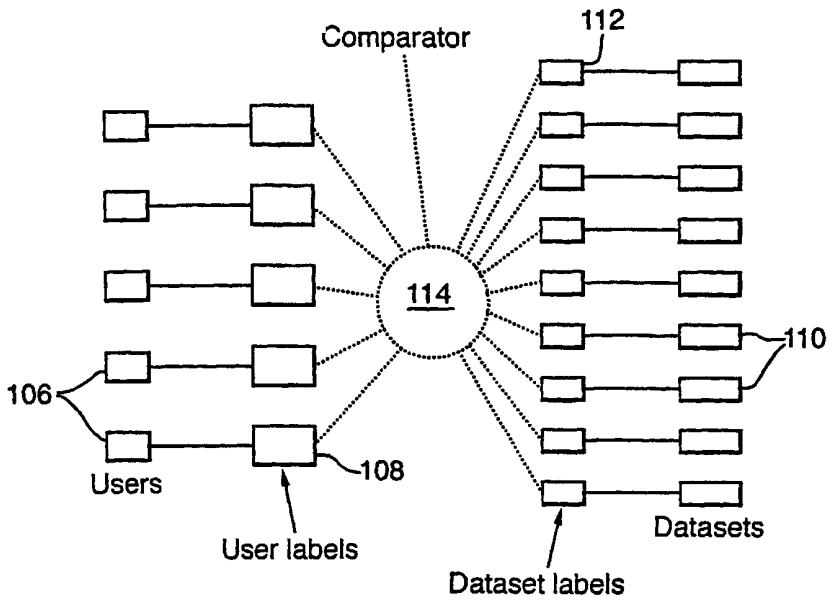
(81) Designated States *(national)*: CA, CN, GB, JP, KR, US.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:
— *With international search report.*

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

(54) Title: METHOD FOR COMPUTER SECURITY



(57) Abstract: A method for computer security to control access to data held on a computer system comprises labelling datasets (110) stored in it with dataset access labels (112) associated with one or more security access levels. A user (106) wishing to gain access to a dataset (110) is allocated a user label (108) indicating the security level of datasets to which the user is to be granted access. The user (106) establishes user identity and the user label (108) is compared with the dataset label (112). The user (106) is allowed access only to datasets (110) with a dataset label (112) indicating a security level equal to or lower than that of the user label (108). This method is particularly relevant to Internet applications to determine whether or not users (106) can access datasets (110) consisting of web pages.

WO 01/25882 A1

METHOD FOR COMPUTER SECURITY

This invention relates to a method for computer security, and also to computer security apparatus, a computer network with security provision, and computer software for computer security.

It is a long felt want to be able to restrict access to data held on a computer system. Various techniques are known which are intended to provide for access to such data to be reserved to those authorised to use it. However, computer hacking (unauthorised access to a computer system) is well known to occur, and a computer system holding data subject to restricted access needs to be as secure as possible: i.e. there should be no security loop holes which can be exploited by potential hackers.

Security is particularly important in systems networked together via the Internet, where there is a problem controlling access of individuals and groups to web pages made available to users on a web server computer.

European Patent Application No. 0 848 314 A1 (corresponding to US Pat No 6,006,228 to McCollum et al) discloses controlling user access to documents by an automated equivalent of conventional paper procedures. Clients are given identity indicators, and both clients and documents are given clearance level indicators: both kinds of indicator are stored on a computer system accessible via the Internet. A client wishing access to a document provides an identity indicator, i.e. a client name, and the computer system looks up and compares the clearance levels of the client and document. If the clearance level of the client is equal to or above that of the document the client is given access. This procedure has the disadvantage that each individual client has to be given a computer identity and a clearance level both requiring entry into a table on a computer system. It is

onerous for large numbers of users, e.g. employees of large organisations or companies such as banks with large numbers of customers.

It is also known to control access to data in the form of web pages via the Internet by means of what is referred to in the prior art as a "secure socket layer" (SSL) (IETF Internet Draft SSL Protocol version 3): the SSL uses public key cryptographic technology. Public key algorithms use one key for encryption and a different key - a "private" key - for decryption; however, the decryption key cannot be calculated from the encryption key (not at least in a reasonable amount of time). The public key can be made public and used by anyone to encrypt a message. Only the person with the associated private key can decrypt the message.

Messages may be encrypted with a private key and decrypted with a public key. This allows any party to authenticate a message which is from the owner of the private key. SSL makes use of this as a method of client identification to authenticate clients: it encrypts a communication with the private key and sends it for decryption with the public key. Successful decryption authenticates a client and may be referred to as a handshake procedure. After this initial handshake, a symmetric session key is generated between the server and client and used to encrypt subsequent communications. After a period of time, another handshake may be done, and a fresh session key generated (to prevent overuse of an individual session key and its consequent exposure to cracking). A client authenticated in this way is allowed to access directories and web pages that require presentation of a valid certificate, e.g. an X.509 Certificate. The X.509 Certificate is a certificate embodied in computer code and is computer readable. It is obtainable by a client from a variety of certifying authorities such as computer system vendors. It contains a public key for the requesting client and other information that serves to identify the client uniquely in accordance with the

standard set by the International Telecommunication Union - ITU-T Recommendation X.509.

The X.509 certificate technique does not provide for varying levels of access: a client seeking data in a web server computer either does or does not gain access to it depending on whether or not the computer finds that the client's X.509 certificate is authenticated for this purpose. There is no provision for access to some data but not others.

As has been said, SSL involves a handshaking procedure allowing a server and client to authenticate one other and to negotiate an encryption algorithm and cryptographic keys before the application protocol transmits or receives its first byte of data. Furthermore, because SSL is an encryption process, like any such it leads to a major performance degradation: it employs symmetric cryptography for data encryption (e.g., DES[DES], RC4[RC4], algorithms), Encryption prevents file compression for telecommunication transmission. In consequence, encrypted files have to be transmitted in uncompressed form, which clearly requires a greater transmission bandwidth, and/or longer transmission time.

The present invention provides a method for computer security to control access to data held on a computer system as requestable datasets characterised in that it includes:

a) allocating computer system users between a plurality of user groups, each user group corresponding to a respective data access category selected from a plurality such categories;

b) associating each dataset with a dataset access category; and

c) giving access to each dataset only to user group members associated with an appropriate data access category for that dataset.

The method of the invention provides the advantage that it removes the need to assign clearance levels to individuals and store their details on a computer system. Merely assigning a data access category to an individual member of a user group (e.g. customer, staff, management) enables all members of that group to gain access to corresponding datasets. The clearance procedure is then simply that of defining an individual to be a user group member.

The user groups and data access categories may have hierarchical levels in which a higher data access category incorporates a or as the case may be each lower data access category, and the method includes allowing access to datasets by members of user groups associated with data access category levels equal to and higher than those to which such datasets correspond.

Each user may be associated with computer-based identifying means such as an X.509 certificate and the method may include the step of determining a user's identity from the identifying means.

The datasets may be web pages and the method may include the step of gaining access to the computer network via the Internet or the World-Wide-Web. Each dataset may be associated with a dataset access category implemented by inserting meta tags in html web page code. A challenge-response exchange regarding user identification may be performed before giving access to a dataset.

In the method of the invention, user may employ a user computer system to gain access to datasets to which access is controlled by an access control computer system having a public key for verifying signed data, characterised in that each user computer system incorporates a private key for signing data and user group identifying means and the dataset access step includes:

a) using the private key to sign test data (e.g. random data) provided by the access control computer system and forwarding the signed data and identifying means to the access control computer system;

b) using the access control computer system to

c) verify the identifying means,

d) verify the user by using the public key to verify the signed data, and

e) determine user group and associated data access category from the identifying means.

The method may involve providing database access to a first kind of user having a user certificate for identification purposes and a second kind of user lacking such certificate.

Data may be maintained on a database computer system, dataset access being given by access control software operated on a separate access control computer system, and a user gaining access to data by means of access request software running on a user computer system separate from the database and access control computer systems. The access control software may be configured with a firewall protecting a database computer system.

Data may be maintained on a plurality of database computer systems and in response to a data request the access control software may determine whether or not corresponding data access is appropriate after relaying the request to a dataset computer system having such data.

Data access categories and the user groups and datasets with which they are associated may be assigned respective numerical values and the step of giving dataset access then involves comparing user group and dataset numerical values to determine whether or not access is to be granted or denied. The data access categories may have different sections each with a section numerical value and

the step of comparing numerical values may comprise comparing section numerical values of corresponding sections of user group and dataset numerical values.

The step of giving access to a dataset may include unencrypted transfer of data from datasets to which access is granted. The method may include the step of running checking/blocking software on a user computer system to screen incoming data for encryption to block unwanted data content.

In another aspect, the invention provides a computer program for controlling operation of a computer system and providing control of access to data held on a computer system as requestable datasets characterised in that the computer program is arranged to:

a) receive data requests from computer system users allocated between a plurality of user groups, each user group corresponding to a respective data access category selected from a plurality of such categories;

b) control access to datasets each of which is associated with a dataset access category; and

c) give access to each dataset only to user group members associated with an appropriate data access category for that dataset.

The user groups and data access categories may have hierarchical levels in which a higher data access category incorporates a or as the case may be each lower data access category, and the computer program may be arranged to allow access to datasets by members of user groups associated with data access category levels equal to and higher than those to which such datasets correspond.

The computer program may be arranged to determine a user's identity from computer-based identifying means such as an X.509 certificate.

The datasets may be web pages, the computer program enabling access to the web pages via the Internet or the World-Wide-Web and identifying dataset access categories in web pages from meta tags in html web page code.

The computer program may be arranged to challenge incoming data requests regarding user identification before giving access to a dataset.

The computer program may be for interacting with a user computer system incorporating a private key for signing data and user group identifying means, and be arranged to:

a) send test data to the user computer system for signature with the private key and return with the identifying means,

b) verify the identifying means,

c) verify the user by using the public key to verify the signed data, and

d) determine user group and associated data access category from the identifying means.

The test data may be random data.

The computer program may be arranged to provide database access to a first kind of user having a user certificate for identification purposes and a second kind of user lacking such certificate. It may be configured with a firewall for a database computer system. It may be arranged to transfer dataset material to appropriate recipients unencrypted.

Data access categories and the user groups and datasets with which they are associated may be assigned respective numerical values and the computer program grants or denies dataset access on the basis of comparison of user group and dataset numerical values.

In a further aspect, the invention provides a network access controller for controlling access to data held on a computer system as requestable datasets the controller being arranged to:

a) receive data requests from computer system users allocated between a plurality of user groups, each user group corresponding to a respective data access category selected from a plurality of such categories;

b) control access to datasets each of which is associated with a dataset access category; and

c) give access to each dataset only to user group members associated with an appropriate data access category for that dataset.

The controller may be adapted to compare numerical values associated with data access categories of datasets and user groups in order to determine whether or not to grant access to data.

The controller may provide database access to a first kind of user having a user certificate for identification purposes and a second kind of user lacking such certificate.

In another aspect, the invention provides a computer network for database access by users allocated between a plurality of user groups and having identifying certificates, the network being arranged to treat each user group as corresponding to a respective data access category selected from a plurality of such categories, and including:

a) an access controller controlling access to a database comprising a plurality of datasets each having an associated dataset access category,

b) means for verifying users,

c) a database of datasets each of which is associated with a dataset access category; and

d) computer software arranged to give access to each dataset only to user group members associated with an appropriate data access category for that dataset.

The network may be an Internet or World-Wide Web network and the database may comprise web pages in which dataset access categories are implemented by insertion of meta tags in web page html code.

There now follows, by way of example only, a description of an embodiment of the invention with reference to the accompanying drawings of which:

Figure 1 is a schematic block diagram of a computer network arranged for security in accordance with the invention;

Figure 2 indicates contents of a user certificate and access control list for use in the access process of the invention;

Figure 3 is a block diagram indicating an access process of the invention for client access to data held on a computer system;

Figure 4 is a flow diagram of the operation of the computer network of Figure 1;

Figure 5 indicates a prior art Internet approach to mapping user identities to each potentially accessible dataset;

Figure 6 indicates user identification by user label, dataset categorisation by dataset label, and label comparison in accordance with the present invention; and

Figure 7 schematically shows comparison between user and dataset labels.

The computer network shown in Figure 1 comprises a computer system, in the form of a web server 2, containing datasets consisting of web pages 4 to which

clients of a web service provider may wish to have access in order to engage in web browsing. The web server 2 is connected via a conventional network link 5 to a further computer system 6 (indicated within chain lines) running a proxy server 8 that maintains an access control list 10. The proxy server 8 controls access of clients to the web server 2. The access control list maintains a set of security clearance levels (e.g. three) expressed as user label entries and associated with client Distinguished Name categories (as will be described in more detail later). The user label entries are associated with degrees of access to web pages which are available to Distinguished Name categories.

A further conventional communications link 14 connects the proxy server 8 to an external client computer 12 (indicated within chain lines) requiring access to the web pages 4. The client computer 12 is running at least two separate software applications: i.e. network access software consisting of a web browser 16, and client proxy software 18 which forwards communications to the proxy server 8 and responds to authentication requests from the proxy server 8. The client computer 12 has stored within it an X.509 certificate previously issued to the client using it, the certificate having been issued by an issuing body which either controls access to the web server 2 or is approved by whoever controls such access. The client software 18 and the X.509 certificate 20 are both necessary for access to the web server 2.

The X.509 certificate includes what is referred to in the computer field as a "Distinguished Name": this name is used in accordance with the invention to identify the client uniquely at least in so far as information access control via the web server 2 is concerned: the Distinguished Name is made up of a hierarchical set of address elements or components in ascending or descending size order, such as for example Country, State and/or County, Town, Organisation, Organisational Unit and Common Name (i.e. personal name). An example of a

Distinguished Name is Britain, Worcestershire, Droitwich, Moneybags Bank, Accounts Department, John Smith.

The present invention makes use of the Distinguished Name principle to avoid the need to give individual clearances to clients and instead use the groupings to which they belong. Rather than associate a particular clearance with each individual client, the AWAC system allows large categories of people to be associated with a clearance. In the above example, anyone with Country:Britain might be given a lowest level clearance, Organisation:Moneybags Bank might be the next level up and Organisational Unit:Accounts Department might receive a top level clearance.

The computer system 6 running the proxy server 8 accommodates data requests from another type of computer 22 (indicated within chain lines) running access software 24 (e.g. a web browser) but lacking other necessary items. Client software 26 and an X.509 certificate 28 are held by the computer system 6 for this purpose and data passes between the computers 6 and 22 through a connection 30.

Each item of data held on the web server 2 has a security level associated with it. The security level is implemented as a data label or document label and is incorporated in the data as what is referred to as a "meta tag": a meta tag is a facility in the World-Wide-Web html (hyper-text mark up) language for adding to a document information which will not be displayed on a visual display unit screen to a client having access to the document. Absence of visibility is not however essential, but it avoids a visible document being adulterated with unwanted material.

An appropriate security level is assigned to each web page 4. In order to gain access to any particular web page, a client for the web service to which access is controlled in accordance with the invention is either has been previously issued

with and identified by an X.509 certificate 20, or be allowed access via the alternative link 30 and using the certificate 28. The components of the distinguished name in the certificate 20 or 28 are compared to entries in the access control list 10. Components of possible distinguished names entered in the access control list have respective security access levels associated with them. For data access to be granted, the components of the distinguished name in the certificate 20 or 28 must correspond to a security access level which is at least as high as the security access level or data label of the web page requested by the client.

There is a plurality of security access levels in the access control list 10, for example in ascending order of restrictedness: unclassified, restricted, secret. A client having permission for access to secret data can access data at all these three security levels; a client having permission to gain access to restricted data can see both restricted and unclassified data but not secret data, whereas a client cleared for access to unclassified data only can see such data but not restricted or secret data. This hierarchical approach is not essential: each type of clearance might only allow access to a respective and single type of information. For example, in a database access application, one might restrict access to personal information to those with the appropriate surname: this corresponds to a code word approach to access control as opposed to the hierarchical scheme. It is also possible to restrict access by a caveat approach, e.g. imposing a restriction such as "Management Only".

Each of the web pages 4 held on the web server 2 is assigned a respective appropriate security level (i.e. unclassified, restricted or secret) by inserting an html meta tag as described earlier, and only clients with clearance to at least that level will be given access to that page.

Use of a meta tag in a web page avoids the need for a computer to hold details of individual documents and their security access levels. Instead an appropriate meta tag is inserted once and for all and then the computer 2 checks it when the document is requested.

The security access level of a particular client is determined by the client's Distinguished Name in the X.509 certificate 20 or that presumed from the certificate 28: e.g. one or a combination of two or more of Country, State and/or County, Town, Organisation, Organisational Unit and Common Name. The proxy server 8 obtains the Distinguished Name identifying the client from the certificate 20 or 28 and determines the associated security access level from the access control list 10. It controls whether or not a requested page held on the web server 2 can be accessed by that particular client in accordance with the client's security access level. Because all communication must pass through the computer 6 running the proxy server 8, it is more secure than systems which provide access control on the same computer system that is providing access to data, e.g. running the web server 2. For example hackers may be able to exploit loopholes in a computer system as a whole, e.g. by means of manipulating operating platform software: they may gain access to data held on web pages and thus by-pass the access control. If the operating software of the database server is well-known, e.g. Windows NT or other widely available commercial software, there may be many who are aware of its potential for loopholes.

The Distinguished Name in the X.509 certificate 20 or 28 is as has been said a series of components or elements giving geographical locations and/or organisational groupings that the client is a member of together with the client's name. Each of these elements, and combinations of them, may be a means of identifying groups of people. Not all elements of a Distinguished Name in the certificate need to be used for a security access level, for example a Distinguished Name may be:-

Country (C) = Great Britain (GB),
Organisation (O) = The Zoo,
Organisational Unit (OU) = Elephants,
Common Name (CN) = Mark.

With this Distinguished Name, Mark is a member of the following four groups:

People from Great Britain.
People from Great Britain who work for 'The Zoo'.
People from Great Britain who work for 'The Zoo' in the Elephants department.
People from Great Britain who work for 'The Zoo' in the Elephants department and are called Mark.

The State and Location elements of Mark's address have not been completed.

Mark therefore has a Distinguished Name "GB/Zoo/Elephants/Mark" in this example. Having details of groups to which Mark belongs, the proxy server 8 checks the access control list 10 to see if there are any user label entries that match Mark's groups. There may be any number of entries that match, and each entry will have associated with it a user label indicating degree to which access is restricted, i.e. level of security (e.g. "Unclassified", "Restricted" or "Secret"). For example, for the groups to which Mark belongs the access control list could have security level user labels of "Unclassified" for "GB", ","Restricted" for "Zoo", "Secret" for "Elephants" and "Mark" unassigned. Each element (GB, Zoo, Elephants or Mark) of the Distinguished Name (GB/Zoo/Elephants/Mark) may map to a security level user label (or to a corresponding element of such a user label). In a user label library or register there need not necessarily be a respective user label for each element of the Distinguished Name, but only for one or more such labels

Referring now also to Figure 2, in which parts described earlier are like referenced, there is shown a representation of the contents of an X.509 certificate 20, access (tick)/no access (thumbs down) indicators appearing on web pages 4a to 4e and an access control list 10. The X.509 certificate 20 is indicated as a scroll but is in fact a string of computer code. The expression "AWAC" is an acronym from "authenticated web access control", a title for the invention. A Briton (C) Joe Bloggs (CN) is a member of the research department (OU) of an organisation (O) AWAC Inc. As indicated by ticks on web pages 4a to 4c, he is cleared to receive anything up to medium. Thumbs down on web pages 4d and 4e indicate that he is not cleared for anything above medium. The access control list 10 contains the four categories C, O, OU, CN together with a security label and security rating. It indicates that all AWAC Inc. employees are cleared to received unclassified material, Research Department employees can receive material up to "low/medium" classification, Mr Bloggs can receive material up to "Medium" classification , and Personnel Department employees can receive material up to "High" classification.

Referring now also to Figure 3, in which parts described earlier are like referenced, a web access procedure indicated generally by 40 is illustrated. Under the control of a client, the web browser 16 generates a web access request 16r and sends it to the client proxy software 18 for transmission at 18r to the proxy server 8 and thence at 8r to the web server 2. These requests are expressed in accordance with the communications protocol http (hypertext transfer protocol).

The web server 2 responds by sending the proxy server 8 an http message 2m incorporating web page material expressed in html as defined earlier. The proxy server 8 then generates a request 8c for the client's X.509 certificate and signature, and sends it with random data to the client proxy 18, which responds at 18c with the certificate and the random data signed with the client's private key. Signature is implemented by the client proxy 18 using a publicly available

"hashing" algorithm to operate on the random data and turn it into a unique string of code for "signature" - i.e. encryption - using the client's private key. As indicated by 8x, the proxy server 8 performs a series of checks on the certificate, firstly whether or not it is time expired and from an acceptable issuer. It then ascertains whether or not the data was signed with the private key by attempting to decrypt it using the public key contained in the X.509 certificate.

If decryption is successful indicating that the data was in fact signed using the client's private key, proxy server 8 checks for user security label entries in the access control list 10 matching those contained in the X.509 certificate. If there are no such entries the proxy server 8 denies access. If one or more such entries are found in the access control list 10, the proxy server 8 performs a security level check on the requested web page, which contains a dataset security label and an associated security level in meta tag form. If the client's Distinguished Name corresponds to an equal or greater security level as compared to (i.e. "dominates") that of the relevant web page label, then the proxy server 8 gives the client access to that page. It produces a response message 8m which incorporates either the web server message 2m or a notification that access is denied depending on whether or not the data was signed using the correct key, and whether or not the appropriate security level was present: the message 8m is passed on at 18m to the web browser 16.

This example of the invention uses a mechanism for requesting and issuing X.509 certificates with their associated public and private keys, which is a trusted process. It is not in fact essential to encrypt and decrypt with such keys in this way but it is often a very useful feature. Moreover, an X.509 certificate is not essential, and it can be replaced by some form of certificate implemented in computer code and incorporating categories of clients which correspond to different degrees of access. The issuer public certificate, i.e. the public certificate of the issuing Certification Authority, is in this example present on the proxy

server computer 6, and the public certificate and associated private key of the client is on the client computer 12. The client's X509 certificate incorporates the signature of the issuing Certification Authority: the proxy server 8 validates this signature using the Certification Authority's public key in its possession. This is how the proxy server 8 verifies that the X509 certificate was issued by an acceptable Certification Authority.

The client's web browser 16 is configured in such a way that it points to the IP address and port number of the client proxy (or client software) 18 and all web file requests go via this proxy. The web browser 16 has a menu option available allowing all file requests originating within it to be sent to an IP address and port number specified by the client. The client proxy 18 forwards the web file request to the proxy server 8, and the proxy server 8 forwards requests to the web server 2. There is a separate software configuration application to specify the IP address and port number of the web server 2 to which the proxy server 8 forwards requests.

The procedure 40 is shown in more detail in Figure 4: here abbreviations have been used in a key 48 and in the remainder of the drawing as indicated in the tables below:

Term	WB	WS	BRSR	SVR	CP
Meaning	Web Browser	Web Server	Browser	Server	Client Proxy

Term	SP	RND	CERT(S).	ACL
Meaning	Server Proxy	Random	Certificate(s)	Access Control List

When a client issues a web page request 50 using the web browser 16, the client software 18 conveys it unaltered at 52 to the proxy server 8 which forwards it unaltered at 54 to the web server 2. After accessing the web pages 4, the web server 2 generates a http response 56: upon receipt of this response, and assuming the requested page exists, at 58 the proxy server 8 generates a string of random data and passes it with a request for the client's certificate to the client software 18.

The client software 18 uses public domain algorithms to apply a digital signature to the random data using its private key as described earlier, and as shown at 60 passes the signed data and a copy of its X.509 certificate from a client certificate store 62 back to the proxy server 8. The X.509 certificate incorporates an expiry date after which it is not valid. The proxy server 8 then performs a number of checks as outlined within a box 64.

The proxy server 8 has a set of certificates from approved issuers of client certificates stored in an issuer certificate store 65. The proxy server 8 checks the issuer certificates to find one that corresponds to the client certificate. Each issuer certificate contains a public key which the proxy server 8 uses at 66 to verify that a client certificate has been correctly signed by one of a number of acceptable issuing bodies. Failure to find an appropriate issuer certificate public key results in the client certificate being invalidated and access being denied. This is necessary to prevent a client manufacturing its own certificate and gaining access to the web server 2.

After finding an appropriate issuer certificate, at 68 the proxy server 8 checks that the client certificate has not expired, and if not, it takes the public key from the client certificate and verifies at 70 and 72 that the random data has been signed correctly by the client and is the same data that the proxy server 8 issued. It then

compares entries in the Access Control List 74 with the security level user labels associated with the groups of the Distinguished Name in the X.509 certificate.

Assuming all of the above checks have passed successfully, at 76 the proxy server 8 compares the security level user label associated with the client entry in the access control list 10 with that stored as a meta tag in the html source code of the requested web page. If the client's security level user label "dominates", i.e. corresponds to an equal or higher clearance level as compared to that of the requested web page, then the web server's original response is conveyed at 78 unaltered to the client software 18 at 80 and then to the web browser 16 at 82 for display on a visual display unit (VDU, not shown). If any of these checks fail, access is denied to the client and a message is returned to the client in web page format stating the reason for access denial. These denials are shown by the boxes 84 to 94 in Figure 4.

Possible reasons for access denial include: an invalid certificate response at 84, 86 or 88 due to an unverified client certificate, certificate expiry date exceeded, or unverified signed data; other possibilities are an incorrect signed data response 90, no matching entry found at 92 in the access control list 10, or absence of security clearance to a sufficiently high level at 94.

The access control list 10 is held in a database with a front end that prevents any alterations being made to the database design: it may take the form outlined in the table below, in which "WWF" indicates World Wild Life Fund.

Country	Organisation	Organisational Unit	Common Name	Security Label	Security Marking
GB				[3-5]	Medium
GB	'The Zoo'			[2-5]	Medium-High
US				[5-5]	Low
US	WWF	Zoo Research		[3-5]	Medium
US	WWF	Zoo Research	Alvin	[1-5]	High

As shown in Figure 1 two paths of access to the proxy server may be provided (via the client software 18 and via the link 30 by-passing the client software 18). Access via the link 30 may be used to allow people to make "anonymous" access to the system or for people who do not have the necessary client software 18 on their computer 22. Of course, it will be realised that if the necessary client software 18 is not running that it will not be possible to verify the identity of the client and that therefore the security method described herein will not be applicable. In such circumstances it might be appropriate for clients accessing the proxy server 18 without the client software 18 to be given a minimum level of access. For instance in the example given above such a client might be given access to unclassified web pages only and prevented from gaining access to those classified restricted or secret. Such a scheme would be realised by providing an entry in the access control list assigning the appropriate security marking someone without a certificate or who is accessing the proxy server 8 anonymously. Alternatively users given anonymous access could be restricted to those cleared to a higher level.

As well as the client software 18 and the web browser 16 it is necessary to run a client software configuration program on the computer 12. This program configures the client software 18 and provides functions such as allowing the IP address and port number of the proxy server 8 to be provided. If a client accesses

the proxy server 8 anonymously with no identification certificate, their browser software 24 will send requests for web pages through client software 26. It may be desirable to have a separate address for different proxy servers, but in the present embodiment the configuration programme has only one.

In addition to the proxy server 8 and the access control list 10, a configuration program is run on the computer system 6 to store in the latter the IP address and port number of the web server 2. The configuration program also establishes a default security environment or default security grading, this being the security level assigned to a data item such as a web page lacking an assigned label in this regard. This may be the highest level of security (so that only users with the highest security clearance can see it) or the lowest.

Either the computer system 6 or the web server 2 has file labeller software which inserts into web pages 4 meta tags of the correct html format and containing appropriate security labels. The file labeller may have utilities allowing any number of pages to be labelled at once for convenience for operators of the computer system 6 tasked with labelling files stored by the web server 2.

As explained earlier clients with a more senior access level (access to secret data) will be given access to less senior access levels (but not visa versa). This is based on the domination theory outlined in mathematical graph theory and may be implemented using a Unified Labelling Scheme ULS wherein a code is assigned to each access level. The codes of clients and data may then be compared using simple mathematical operations such as NOT and AND to determine whether or not a user is entitled to access a particular access level.

The data from the web pages 4 is transmitted over telecommunication lines in compressed format (unencrypted data compressed). It may be compressed by the web server 2. If the proxy server 8 is close the web server 2, it may compress the

data or another computer may be used. The data received by the client's server 12 is decompressed before it is displayed for viewing (and/or storage by the client). Alternatively, the data may be stored compressed at the web server 2.

Referring now to Figures 5 and 6, these illustrate respectively a prior art security method (which might use SSL) and that of the present invention to permit comparison. Figure 5 shows datasets such as 100 and users such as 102: each dataset 100 must contain a list of all users 102 permitted to have access to it. This corresponds to the prior art requiring a respective virtual connection such as 104 between each piece of data and each user allowed access to it. For illustrational convenience Figure 5 does not show all possible connections 104, but two users 102 are shown connected to all datasets 100.

Figure 6 illustrates the simplification provided by the present invention. Users such as 106 are associated with respective security level user labels such as 108 and datasets such as 110 with respective dataset labels 112. When a user 106 requests access to a particular dataset 110, a comparison process 114 compares the associated user label 108 and dataset label 112 to determine whether or not access will be allowed.

It is not necessary for every user 106 to be given an individual user label in the access control list 10. Part or all of a user's Distinguished Name may itself provide clearance to a predetermined level (e.g. restricted or unclassified), with no user-specific clearance being specified. The X.509 Certificate is user specific and must be applied for, but the access control list 10 is a "permissions table" and entry of a client group or type on this is not required to be user-specific: a whole category or categories of users can be assigned the same clearance. This is particularly useful in commercial areas such as financial services where it might be desired to give all customers access to certain facilities such as market information without submitting them all to a clearance procedure. This would

require them merely to be entered into the group "customer". Similarly for employees there might be "staff" and "management" groupings corresponding to different clearance levels.

Figure 7 illustrates one embodiment of the comparison process 114. It employs labels composed according to what is known as a Unified Labelling Scheme. A human-readable label may be composed of many different types of marking. The Unified Labelling Scheme takes these different kind of markings, in the form of hierarchies, caveats and category markings and represents them as a single computer-readable bit string.

The computer-readable bit strings are compared in a logical operation to determine whether or one label dominates the other: in this connection "dominates" means "corresponds to a more restricted or higher level"; for example, a 'Top Secret' label would dominate a 'Secret' label. The bit string is represented within software by pairs of numbers, for example, [1-5] representing Top Secret and [5-5] representing Unclassified as indicated in an earlier table.

In Figure 7 human readable markings U1 to U4 at 108 indicate user labels and D1 to D4 at 110 indicate dataset labels: user labels 108 and dataset labels 110 are translated to associated computer readable labels at 116 (user) and 118 (dataset) respectively. As indicated by a bracket 120, computer readable user labels are combined to form a complex user label set, and a bracket 122 indicates computer readable dataset labels combined to form a complex dataset label set. These two complex label sets are compared at 124 to determine whether or not access to data is to be granted (as previously described).

It will be appreciated that labelling of client groupings and web page or other data items are two significant aspects of the invention. Comparison of these two quantities forms the basis of the access control decision, and provides

authenticated web access control. This approach does not require certification authorities and directory servers, and can be used by any web browser and web server without altering their functionality. Control of the network access control computer 6 can lie entirely in the hands of an organisation or individual who can control contents of user labels and dataset labels, and the access control list 10. This approach also provides scalability in its use of grouping of users by elements of Distinguished Name corresponding to user labels and mapping this to a particular security rating or dataset label. Groups of clients can have their access rights determined by the elements of the Distinguished Name in their user label. The X.509 Certificate, or other user-identification certificate, could form the basis of some of the sub-label regions within the user label for each user.

In the specific example given above there are four server-side applications (server proxy 8, server proxy configuration, permissions programme that sets the access control list 10, and file labeller that inserts ULS labels into html files on the web server 2 and two client-side applications (client proxy 18 and client proxy configuration programme). The file labeller may allow multiple selections of web pages (or other datasets) so that a security administrator can easily label many pages at a time.

A simple permissions programme maps elements of Distinguished Names to security labels using a database table to store the values, with a front end that prevents unauthorised alterations. The server proxy configuration writes information to a system registry which is part of the operating system (e.g. Microsoft Windows) used by the access control proxy server 8. It also sets the default security environment which applies an administrator defined security label controlling what happens if a requested web page does not contain its own security label.

The access control proxy server 8 passes web page requests onto the web server 2 and verifies using public/private keys that the random data from the client's proxy server 18 has been signed correctly. If a client is denied access to a page, the access control proxy server 8 may inform the client of the reason why.

The client proxy configuration program 18 on the client's computer system 12 writes the Internet protocol (IP) address and port number of the access control proxy server 8 to which the client server connects, to the computer system registry. It also writes in the computer system registry the IP address and port number of the World-Wide-Web proxy), which will allow the client to use the Internet in the normal fashion. It may also allow the client to specify which particular certificate would be used for a particular attempt to access a dataset via the access control server.

The client proxy 18 is a software application running on the client's computer that receives a web page request from the client web browser 16 and passes it to the access control proxy server 8 or the World-Wide-Web proxy. If the request has been sent to the access control proxy server 8 the client proxy 18 will in return receive a request for the X.509 Certificate and some random data. It will then sign the random data with its private key and send the data and the Certificate back to the access control server.

In practice, the owner of a database may have a database server and an access control server under their control, possibly on their property. The owner would keep and maintain a user label database and a dataset label database. A client of the owner for the database has a client server (e.g. a web browser), together with software providing a client proxy and client configuration proxy, and possibly provided by the owner. This software may be provided on a machine-readable data carrier (e.g. magnetic or optical disc, a tape, EPROM/ROM etc.) or it may be

provided electronically (e.g. via a telecommunication link as an electrical signal or an e.m. signal).

It will be appreciated that any aspect of the present invention can be used in conjunction with any other aspect, and that preferred features of any aspect may also be applicable to other aspects of the invention.

The authenticated web access control system of the invention is characterised by ease of maintenance and update. A maintenance manager for a prior art website access control system has to alter allowable access identities on each web page to remove or add an allowable user, which can be very time-consuming for a large number of web pages. Using the present invention, a manager simply adds a new user label, or deletes an existing user label from the directory of user labels (or breaks the correlation between an identified user and associated specific user label).

Similarly, if an entire category of web pages were to have its security access level changed (for example because a secret project had become public or was to be made public), the maintenance manager can change the labels for those web pages to give them a lower security value using the file labelling application.

The present invention is particularly beneficial in large systems with many users and/or many potentially accessible datasets. There may be hundreds or thousands of permissible users, or more. There may be thousands, tens of thousands or hundreds of thousands (or more) of datasets or web pages potentially accessible. There might be more than one secure web server (database servers) on the network. The access control server may have different addresses for different web servers and be adapted to address the appropriate web server for a request for a particular dataset (web page).

The method of the invention described above can clearly be carried out by an appropriate computer program on a carrier medium and running on a conventional computer system. Such a program is straightforward for a skilled programmer to implement without requiring invention because it involves well known computational procedures. Such a program and system will therefore not be described further.

CLAIMS

1. A method for computer security to control access to data held on a computer system as requestable datasets characterised in that it includes:
   a) allocating computer system users between a plurality of user groups, each user group corresponding to a respective data access category selected from a plurality of such categories;
   b) associating each dataset with a dataset access category; and
   c) giving access to each dataset only to user group members associated with an appropriate data access category for that dataset.

2. A method according to Claim 1 characterised in that the user groups and data access categories have hierarchical levels in which a higher data access category incorporates a or as the case may be each lower data access category, and the method includes allowing access to datasets by members of user groups associated with data access category levels equal to and higher than those to which such datasets correspond.

3. A method according to Claim 1 or 2 characterised in that each user is associated with a computer-based identifying means and the method includes the step of determining a user's identity from the identifying means.

4. A method according to Claim 3 characterised in that the computer-based identifying means is an X.509 certificate.

5. A method according to Claim 1, 2, 3 or 4 characterised in that the datasets are web pages and the method includes the step of gaining access to the computer network via the Internet or the World-Wide-Web.

6. A method according to Claim 1 characterised in that the datasets are web pages and the step of associating each dataset with a dataset access category comprises inserting meta tags in html web page code.

7. A method according to Claim 1 or 4 characterised in that it includes the step of performing a challenge-response exchange regarding user identification before the step of giving access to a dataset.

8. A method according to Claim 1 in which a user employs a user computer system to gain access to datasets to which access is controlled by an access control computer system having a public key for verifying signed data, characterised in that each user computer system incorporates a private key for signing data and user group identifying means and the dataset access step includes:

a) using the private key to sign test data provided by the access control computer system and forwarding the signed data and identifying means to the access control computer system;

b) using the access control computer system to

i) verify the identifying means,

ii) verify the user by using the public key to verify the signed data, and

iii) determine user group and associated data access category from the identifying means.

9. A method according to Claim 8 characterised in that the test data is random data.

10. A method according to Claim 1 characterised in that it includes providing database access to a first kind of user having a user certificate for identification purposes and a second kind of user lacking such certificate.

11. A method according to Claim 1 characterised in that data is maintained on a database computer system, dataset access is given by access control software is operated on a separate access control computer system, and a user gains access to data by means of access request software running on a user computer system separate from the database and access control computer systems.

12. A method according to Claim 11 characterised in that the access control software is configured with a firewall protecting a database computer system.

13. A method according to Claim 11 characterised in that data is maintained on a plurality of database computer systems and in response to a data request the access control software determines whether or not corresponding data access is appropriate after relaying the request to a dataset computer system having such data.

14. A method according to Claim 1 characterised in that data access categories and the user groups and datasets with which they are associated are assigned respective numerical values and the step of giving dataset access involves comparing user group and dataset numerical values to determine whether or not access is to be granted or denied.

15. A method according to Claim 14 characterised in that the data access categories have different sections each with a section numerical value and the step of comparing numerical values comprises comparing section numerical values of corresponding sections of user group and dataset numerical values.

16. A method according to Claim 14 characterised in that access to a dataset is provided only if all section comparisons are satisfied.

17. A method according to Claim 1 characterised in that the step of giving access to a dataset includes unencrypted transfer of data from datasets to which access is granted.

18. A method according to Claim 16 wherein a user has a user computer system characterised in that the method includes the step of running checking/blocking software on the user computer system to screen incoming data for encryption to block unwanted data content.

19. A computer program for controlling operation of a computer system and providing control of access to data held on a computer system as requestable datasets characterised in that the computer program is arranged to:

a) receive data requests from computer system users allocated between a plurality of user groups, each user group corresponding to a respective data access category selected from a plurality of such categories;

b) control access to datasets each of which is associated with a dataset access category; and

c) give access to each dataset only to user group members associated with an appropriate data access category for that dataset.

20. A computer program according to Claim 19 characterised in that the user groups and data access categories have hierarchical levels in which a higher data access category incorporates a or as the case may be each lower data access category, and the computer program is arranged to allow access to datasets by members of user groups associated with data access category levels equal to and higher than those to which such datasets correspond.

21. A computer program according to Claim 19 characterised in that it is arranged to determine a user's identity from computer-based identifying means.

22. A computer program according to Claim 21 characterised in that the computer-based identifying means is an X.509 certificate.

23. A computer program according to Claim 19 characterised in that the datasets are web pages and the computer program enables access to the web pages via the Internet or the World-Wide-Web.

24. A computer program according to Claim 19 characterised in that the datasets are web pages and the computer program is arranged to identify dataset access categories in web pages from meta tags in html web page code.

25. A computer program according to Claim 19 characterised in that it is arranged to challenge incoming data requests regarding user identification before giving access to a dataset.

26. A computer program according to Claim 19 for interacting with a user computer system incorporating a private key for signing data and user group identifying means, the computer program being arranged to:

a) send test data to the user computer system for signature with the private key and return with the identifying means,

b) verify the identifying means,

c) verify the user by using the public key to verify the signed data, and

d) determine user group and associated data access category from the identifying means.

27. A computer program according to Claim 26 characterised in that the test data is random data.

28. A computer program according to Claim 19 characterised in that it is arranged to provide database access to a first kind of user having a user certificate for identification purposes and a second kind of user lacking such certificate.

29. A computer program according to Claim 19 characterised in that it is arranged to provide a firewall for a database computer system.

30. A computer program according to Claim 19 characterised in that data access categories and the user groups and datasets with which they are associated are assigned respective numerical values and the computer program grants or denies dataset access on the basis of comparison of user group and dataset numerical values.

31. A computer program according to Claim 19 characterised in that it is arranged to transfer dataset material to appropriate recipients unencrypted.

32. A network access controller for controlling access to data held on a computer system as requestable datasets characterised in that the controller is arranged to:

a) receive data requests from computer system users allocated between a plurality of user groups, each user group corresponding to a respective data access category selected from a plurality of such categories;

b) control access to datasets each of which is associated with a dataset access category; and

c) give access to each dataset only to user group members associated with an appropriate data access category for that dataset.

33. A controller according to Claim 32 characterised in that it is adapted to compare numerical values associated with data access categories of datasets and user groups in order to determine whether or not to grant access to data.

34. A controller according to Claim 32 characterised in that it is arranged to provide database access to a first kind of user having a user certificate for identification purposes and a second kind of user lacking such certificate.

35. A computer network for database access by users allocated between a plurality of user groups and having identifying certificates, characterised in that it is arranged to treat each user group as corresponding to a respective data access category selected from a plurality of such categories, and it includes:

a) an access controller controlling access to a database comprising a plurality of datasets each having an associated dataset access category,

b) means for verifying users,

c) a database of datasets each of which is associated with a dataset access category; and

d) computer software arranged to give access to each dataset only to user group members associated with an appropriate data access category for that dataset.

36. A network according to Claim 35 characterised in the database comprises web pages in which dataset access categories are implemented by insertion of meta tags in web page html code.

37. A network according to Claim 35 characterised in that it is an Internet or World-Wide Web network.

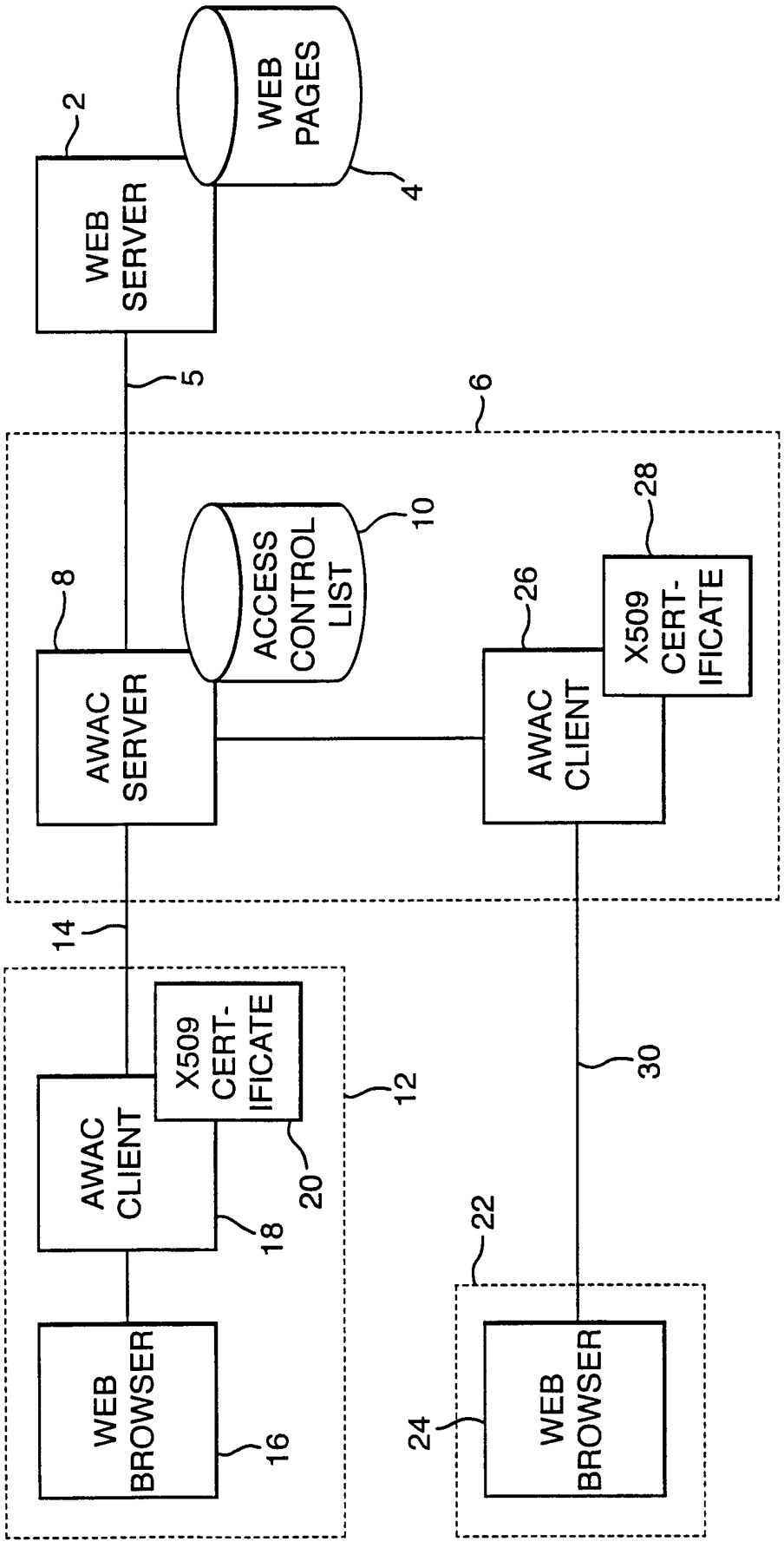
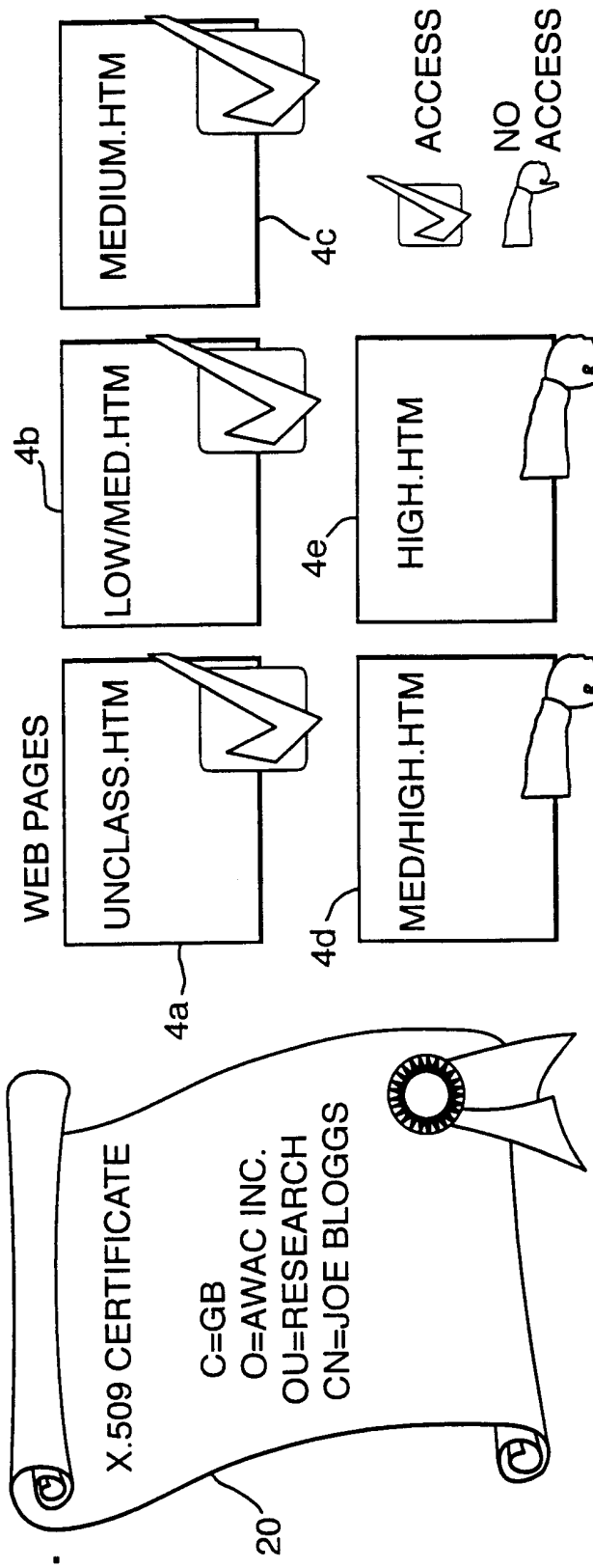

WEB PAGES
4
2
WEB SERVER
5
6
ACCESS CONTROL LIST
10
28
8
26
X509 CERT-IFICATE
AWAC SERVER
AWAC CLIENT
14
X509 CERT-IFICATE
12
30
AWAC CLIENT
20
18
22
WEB BROWSER
16
24
WEB BROWSER

Fig.1.

Fig.2.



ACCESS CONTROL LIST — 10

	C	O	OU	CN	SECURITY LABEL	SECURITY RATING
*	GB	AWAC INC.			345T4	UNCLASS
	GB	AWAC INC.	PERSONNEL		243534	HIGH
*	GB	AWAC INC.	RESEARCH		54534	LOW/MED
*	GB	AWAC INC.	RESEARCH	JOE BLOGGS	45433	MEDIUM

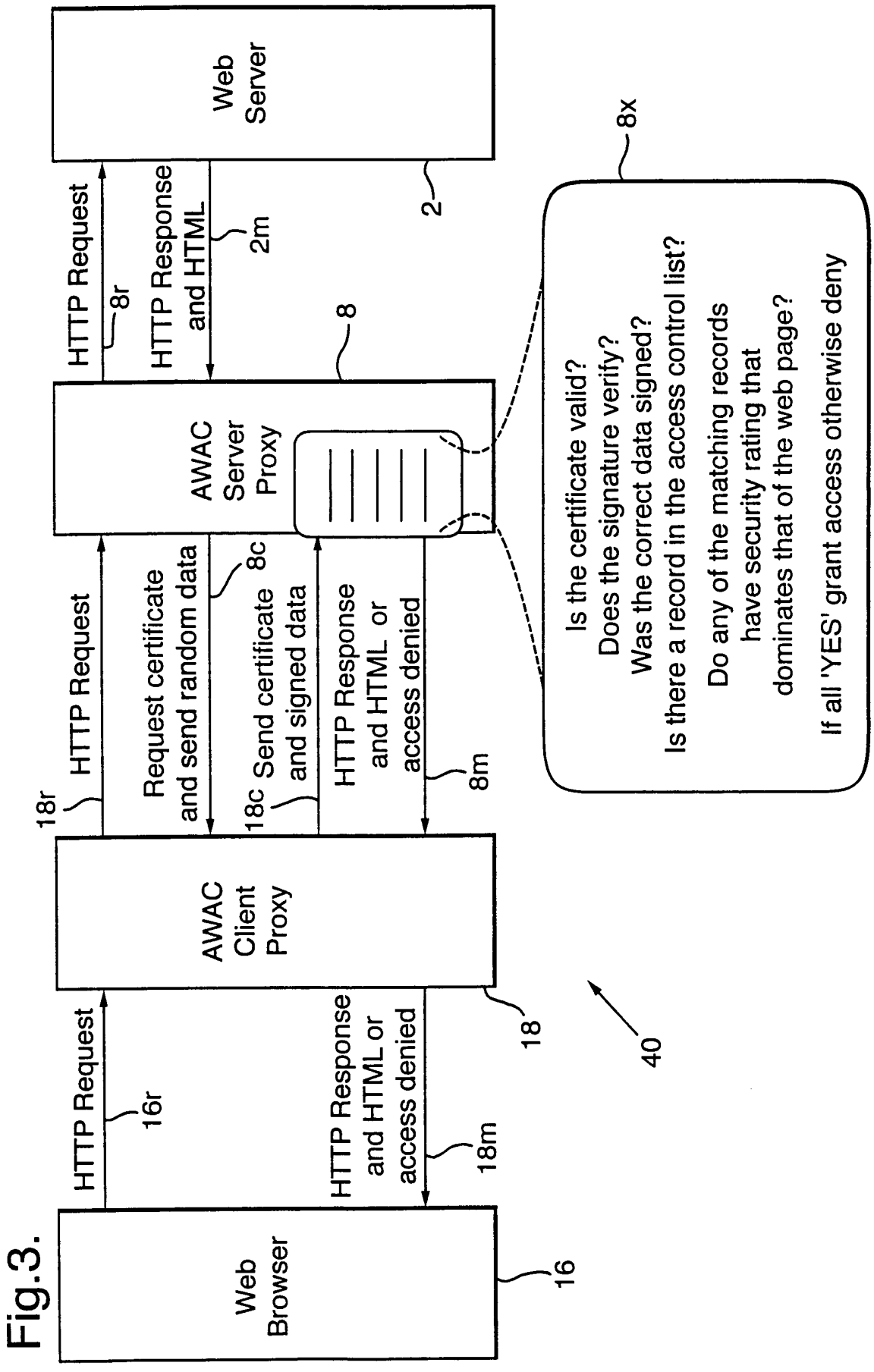

Web Server
2
HTTP Request
8r
HTTP Response and HTML
2m
8
AWAC Server Proxy
8x
Is the certificate valid?
Does the signature verify?
Was the correct data signed?
Is there a record in the access control list?
Do any of the matching records have security rating that dominates that of the web page?
If all 'YES' grant access otherwise deny
HTTP Request
18r
Request certificate and send random data
8c
Send certificate and signed data
18c
HTTP Response and HTML or access denied
8m
AWAC Client Proxy
18
40
HTTP Request
16r
HTTP Response and HTML or access denied
18m
Web Browser
16

Fig.3.

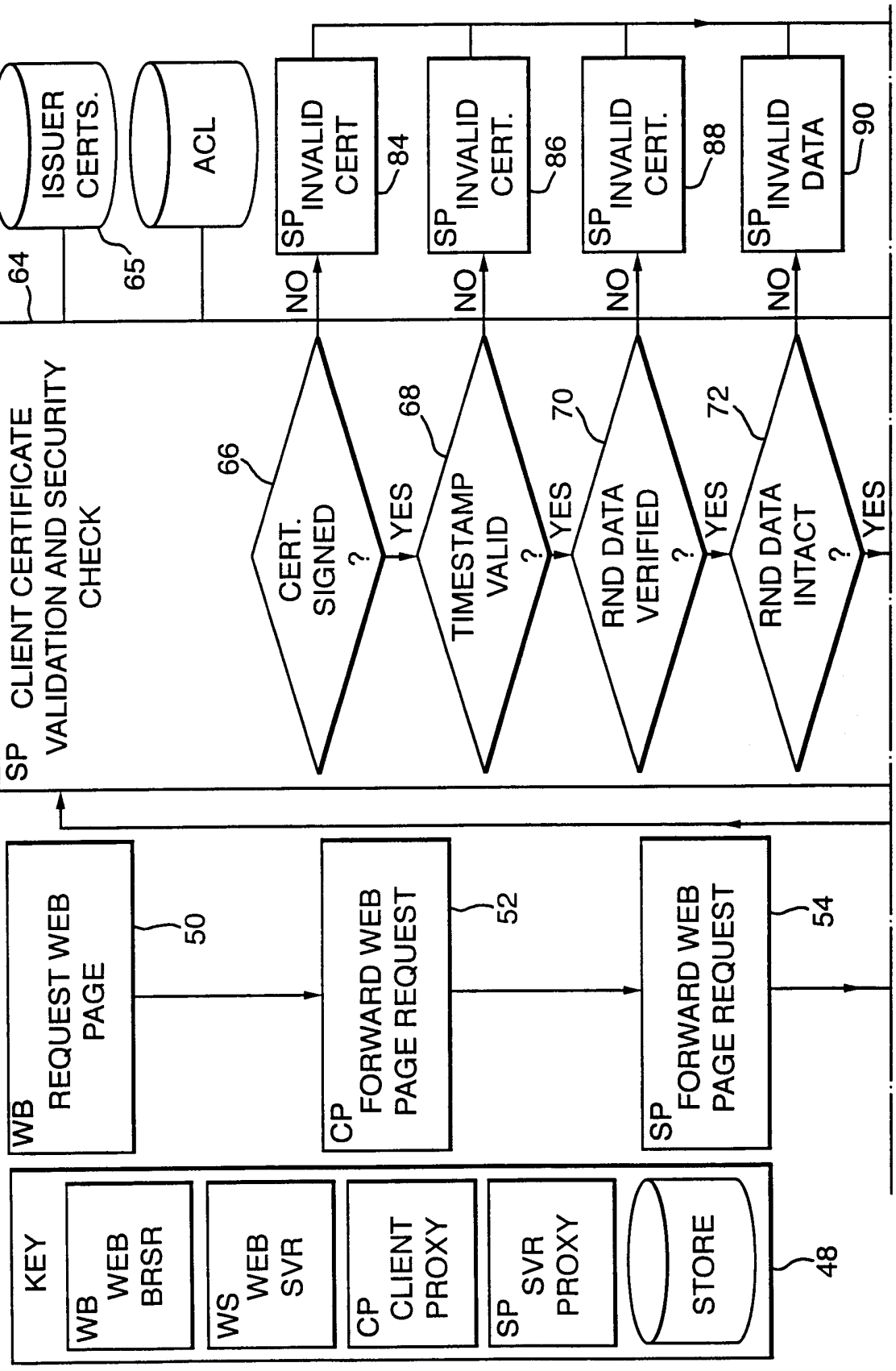

KEY
WB WEB BRSR
WS WEB SVR
CP CLIENT PROXY
SP SVR PROXY
STORE
48
WB REQUEST WEB PAGE
50
CP FORWARD WEB PAGE REQUEST
52
SP FORWARD WEB PAGE REQUEST
54
SP CLIENT CERTIFICATE VALIDATION AND SECURITY CHECK
64
ISSUER CERTS.
65
ACL
CERT. SIGNED ?
66
TIMESTAMP VALID ?
68
RND DATA VERIFIED ?
70
RND DATA INTACT ?
72
YES
NO
SP INVALID CERT
84
SP INVALID CERT.
86
SP INVALID CERT.
88
SP INVALID DATA
90

Fig.4.

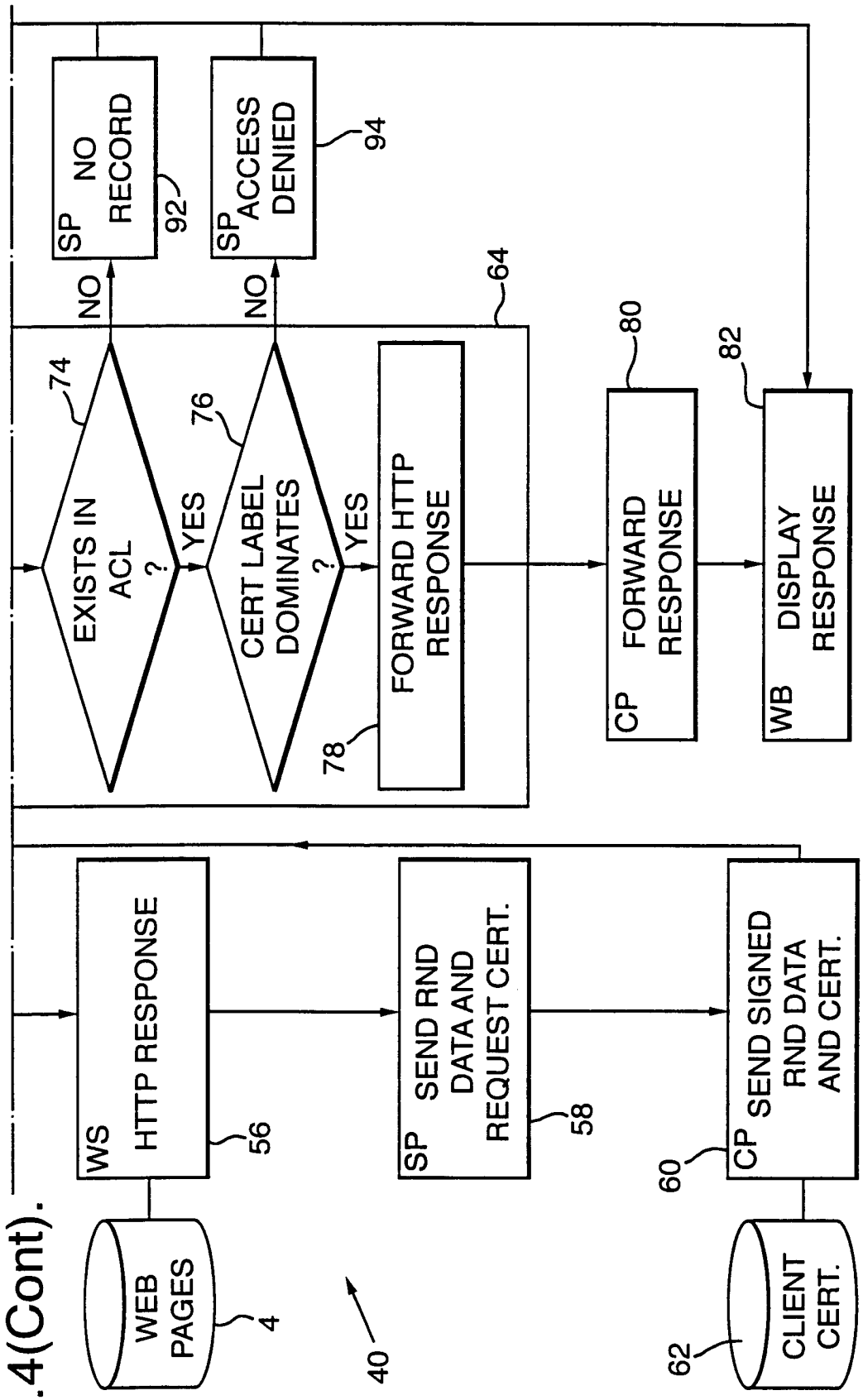

WS HTTP RESPONSE
56
SP SEND RND DATA AND REQUEST CERT.
58
CP SEND SIGNED RND DATA AND CERT.
60
WEB PAGES
4
CLIENT CERT.
62
40
64
EXISTS IN ACL ?
74
YES
CERT LABEL DOMINATES ?
76
YES
FORWARD HTTP RESPONSE
78
NO
NO
SP NO RECORD
92
SP ACCESS DENIED
94
CP FORWARD RESPONSE
80
WB DISPLAY RESPONSE
82

Fig.4(Cont).

Fig.5.

PRIOR ART

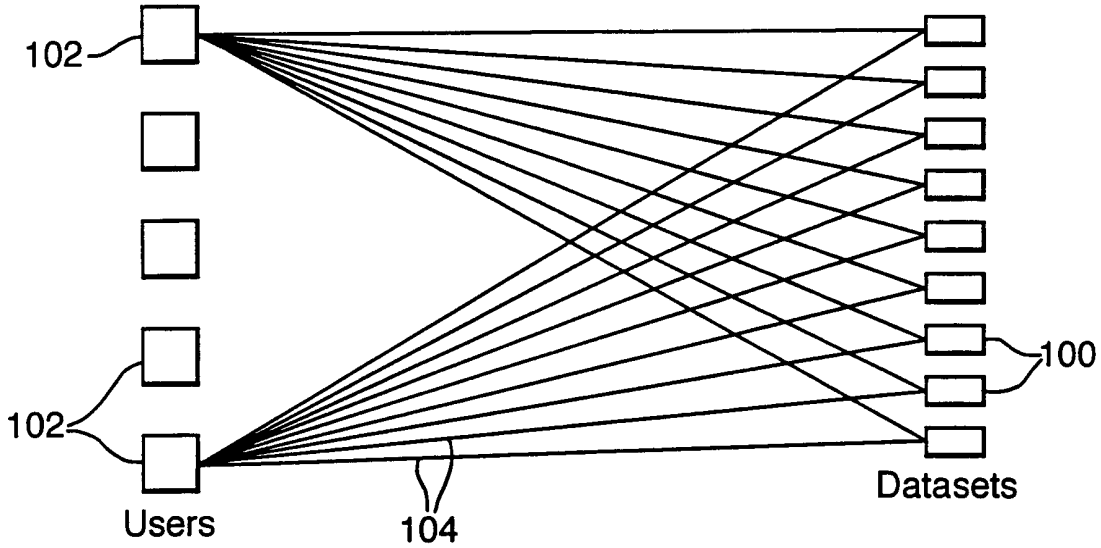


Fig.6.

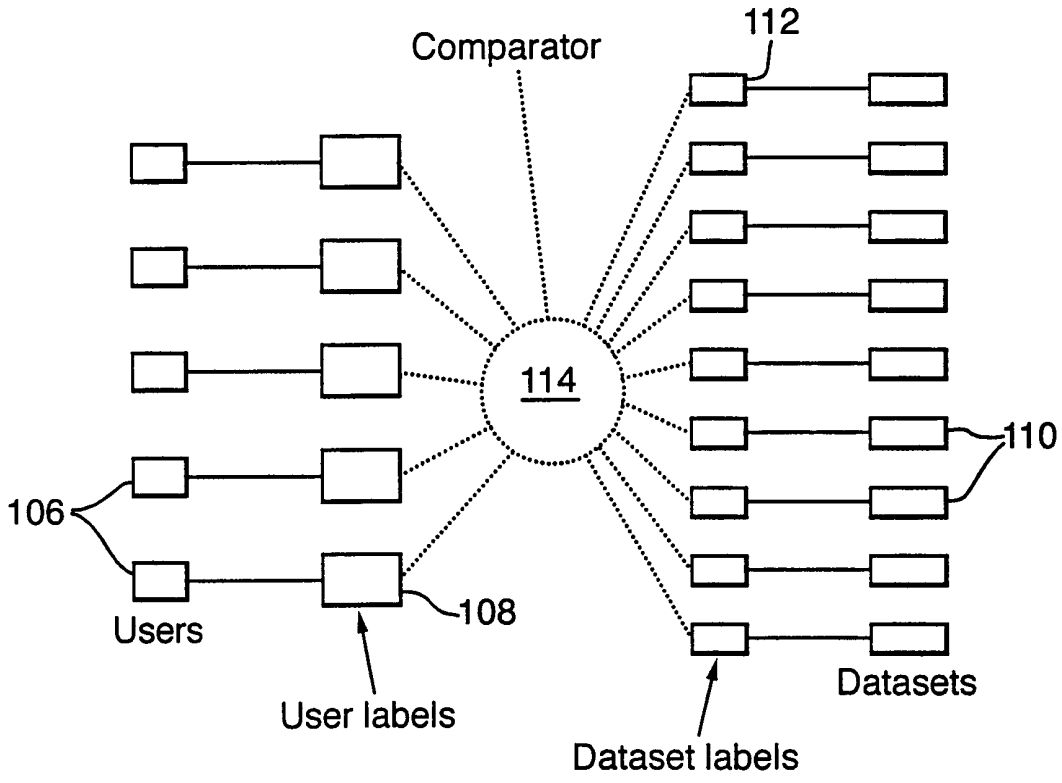
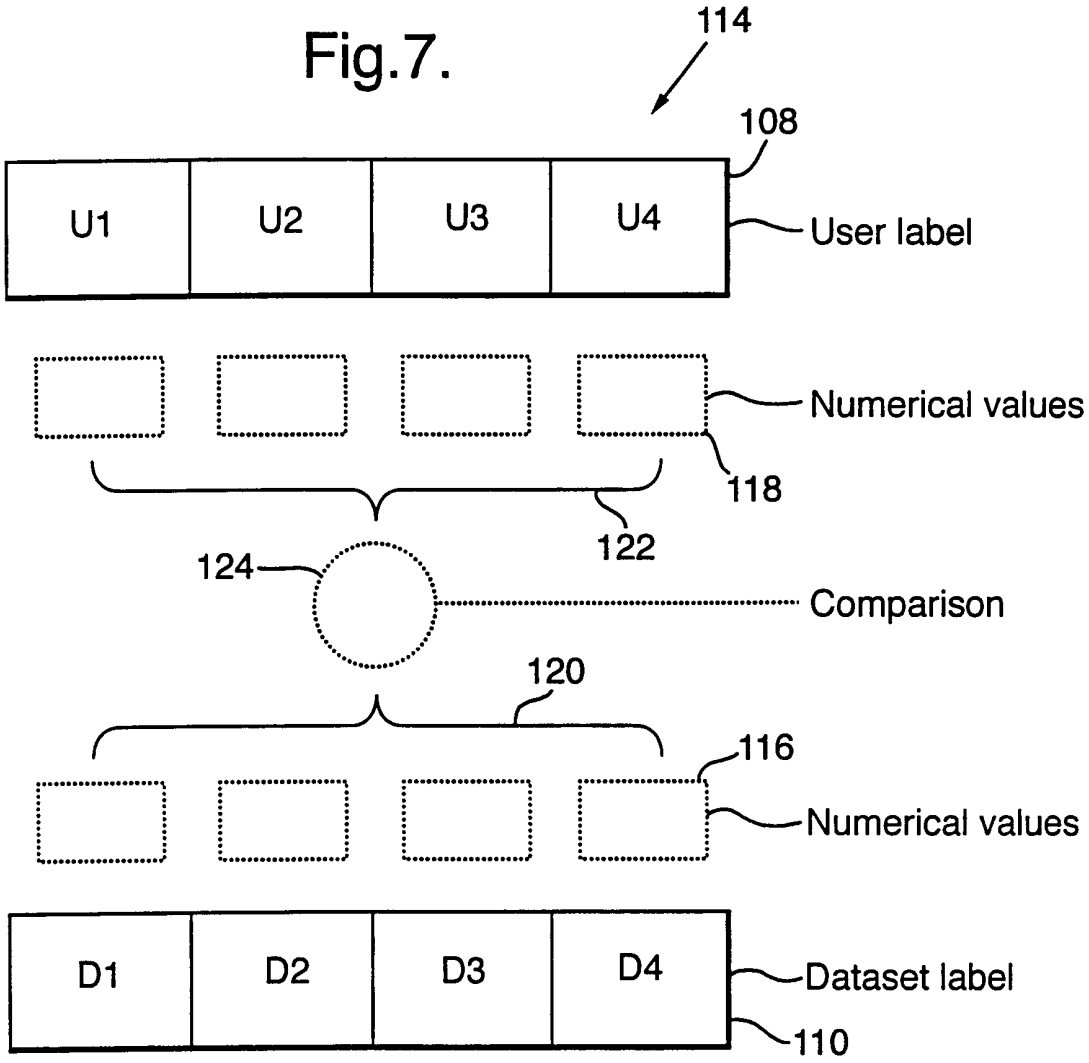


Fig.7.


114
108
U1
U2
U3
U4
User label
Numerical values
118
122
124
Comparison
120
116
Numerical values
D1
D2
D3
D4
Dataset label
110

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/GB 00/03620

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F1/00 G06F12/14

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F H04L

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 17286 A (BIGGE PETER ;TELIA AB (SE)) 6 June 1996 (1996-06-06) abstract; figures 1,2,4 page 1, line 22 -page 3, line 27 page 4, line 1 - line 24 page 5, line 27 -page 6, line 19 page 8, line 16 -page 9, line 35	1,3,5-7, 10,17, 19,21, 23-25, 28,31
Y		2,4, 11-16, 20,22, 29,30, 32-37

-/--

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

[X] Patent family members are listed in annex.

° Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 December 2000	03/01/2001

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

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interna al Application No
PCT/GB 00/03620

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 407 060 A (AMERICAN TELEPHONE & TELEGRAPH) 9 January 1991 (1991-01-09) the whole document	1,19
Y		2,14-16, 20,30
Y	DAVIS J ET AL: "AN IMPLEMENTATION OF MLS ON NETWORK OF WORKSTATIONS USING X500/509" PHOENIX/TEMPE, FEB. 5 - 7, 1997,NEW YORK, IEEE,US, 5 February 1997 (1997-02-05), pages 546-553, XP000753724 ISBN: 0-7803-3874-X the whole document	4,11-13, 22,29, 30,32-37
A		1,8,10, 17-19, 26,28,31
A	US 5 940 591 A (BOYLE JOHN M ET AL) 17 August 1999 (1999-08-17) the whole document	32-34

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Interna al Application No
PCT/GB 00/03620

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9617286 A	06-06-1996	EP 0795151 A	17-09-1997
		SE 9404157 A	30-05-1996
EP 0407060 A	09-01-1991	CA 2018319 A,C	31-12-1990
		DE 69029880 D	20-03-1997
		DE 69029880 T	21-08-1997
		JP 3041535 A	22-02-1991
US 5940591 A	17-08-1999	US 5577209 A	19-11-1996