

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
20 March 2008 (20.03.2008)

PCT

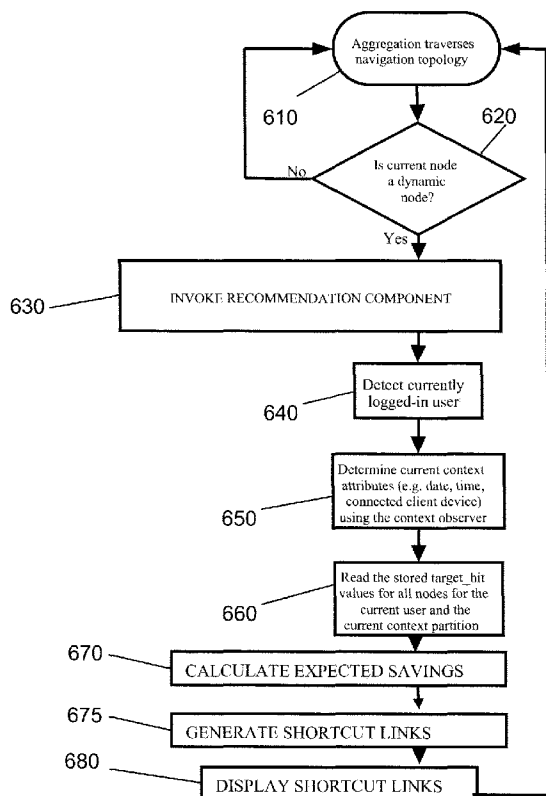
(10) International Publication Number
WO 2008/031646 A1

- (51) International Patent Classification:
H04L 29/08 (2006.01) G06F 17/30 (2006.01)
- (21) International Application Number:
PCT/EP2007/056475
- (22) International Filing Date: 28 June 2007 (28.06.2007)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
06120465.7 11 September 2006 (11.09.2006) EP
- (71) Applicant (for all designated States except US): **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, NY 10504 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **NAUERZ, Andreas** [DE/DE]; Belchenstr. 2, 71032 Böblingen (DE). **LIESCHE, Stefan** [DE/DE]; Gabriele-Muenter-Hof 17, 71034 Böblingen (DE).

- (74) Agent: **KLEIN, Hans-Jörg**; 70548 Stuttgart (DE).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

(54) Title: METHOD FOR RECOMMENDING SHORT-CUTS IN ORDER TO EASE AND FASTEN NAVIGATING WITHIN NAVIGATION TOPOLOGY



(57) Abstract: The present invention relates to the field of network computing, and in particular to method and system for designing a Web Portal comprising a hierarchical structure of portal pages and portlets for accessing Web contents accessible via the Portal. For offering improved web application navigation it is proposed to perform the steps of: a) during navigation of a user in the topology of said web portal automatically sensing (520, 525) information about the navigational behavior of said user for each page comprised of said navigation, b) deriving (530) page hits and page target hits from said sensed information, wherein page hits are defined for pages clicked for reaching another page, and page target hits are defined for pages a user clicks in order to perform a page-specific task or to receive some page-specific information, c) using a predefined metric based on said sensed information for generating a utility ranking for the web pages comprised of the user navigation, d) tracking (620) the current navigational position during said user's navigation, e) calculating for said current position and a plurality of target pages a probability measure to navigate to a respective one of said target pages, wherein said target pages are obtained by said utility ranking, f) calculating (670) an expected saving factor for each target page based on said target page utility and on the navigational distance between said current page and said target page, g) defining (675) at least one shortcut including the respective direct link to at least one of the target pages having a high expected saving factor, h) displaying (680) said short cut as a clickable link to said user.

WO 2008/031646 A1

- 1 -

D E S C R I P T I O N

**Method For Recommending Short-Cuts In Order To Ease And Fasten
Navigating Within Navigation Topology**1. BACKGROUND OF THE INVENTION1.1. FIELD OF THE INVENTION

The present invention relates to the field of network computing, and in particular to method and system for designing a Web Portal or Enterprise Portal comprising a hierarchical structure of portal pages and portlets for accessing Web contents or Enterprise contents accessible via the Portal.

1.2. DESCRIPTION AND DISADVANTAGES OF PRIOR ART

Figure 1 gives a schematic system view on a Portal server implementing such prior art Web Portal.

A prior art Portal as e.g., represented by above IBM WebSphere Portal or by Jetspeed2 Enterprise Portal (www.Portals.apache.org/jetspeed-2/Portal-design.html) is built by a complex functionality implemented on a network server - for example a Web server 100, the most important elements of which are logic components for user authentication 105, state handling 110, aggregation 115 of fragments, a plurality of Portlets 120 - further described below- provided in respective pages 125 with a respective plurality of APIs 130 to a respective Portlet container software 135 for setting them into the common Web page context, and some Portal storage resources 140. The logic components are operatively connected

- 2 -

such that data can be exchanged between single components as required. This is roughly depicted in figure 1.

In more detail, a Portal engine of the Web server in figure 1 implements an aggregation of Portlets 120 based on the underlying Portal model 150 and Portal information such as security settings, user roles, customization settings, and device capabilities. Within the rendered page, the Portal automatically generates the appropriate set of navigation elements based on the Portal model. The Portal engine invokes Portlets during the aggregation as required and when required and uses caching to reduce the number of requests made to Portlets. The prior art IBM WebSphere Portal employs open standards such as the Java Portlet API (application programming interface). It also supports the use of a remote Portlet via the WSRP standard.

The Portlet container 135 is a single control component competent for all Portlets 120, which may control the execution of code residing in each of these Portlets. It provides the runtime environment for the Portlets and facilities for event handling, inter-Portlet messaging, and access to Portlet instance and configuration data, among others. The Portal resources 140 are in particular the Portlets 120 themselves and the pages 125, on which they are aggregated in form of an aggregation of fragments. A Portal database 128 stores the portlet description, this is in detail the portlet description featuring some attributes like portlet name, portlet description, portlet title, portlet short title, and keywords; the portlet interaction interface description, which is often stored in form of WSDL documents. The Portal database also stores the Portal content structure, i.e. the hierarchical structure of portal pages - which may again contain nested pages - and portlets. This data is stored in

- 3 -

the database 128 in an adequate representation based on prior art techniques like relational tables.

The before-mentioned aggregation logic 115 includes all steps that are required to assemble a page. Typically, these steps are to load a content structure from storage, to traverse it and to call the instances referenced in the structure in order to obtain their output, which is assembled to a single page.

The content structure may be defined through e.g. Portlet Customization by the administrator.

With specific focus now to the present invention when web applications are visited by a web user, a user is usually displayed a navigation menu that provides some means to access underlying content. A navigation menu is usually structured in a tree-like topology, and users are forced to traverse the tree in order to reach a node matching the content the user is interested in. Specifically, web portals are equipped with such navigation menus which must be used to navigate through all of the contents the web portal provides.

As a matter of fact, not every user is interested in the same content, and hence the structure which is given on a given portal may satisfy the needs of a certain user group, but for many individual users the given topology does not satisfy their needs. Prior art allows to blend-out certain nodes of the topology, i. e. blend-out a page or parts of a page, basically some portlets. This, however is a very inflexible means to better transverse the navigation tree.

1.3. OBJECTIVES OF THE INVENTION

It is thus an objective of the present invention to provide

- 4 -

method and system offering improved web portal navigation.

2. SUMMARY AND ADVANTAGES OF THE INVENTION

This objective of the invention is achieved by the features stated in enclosed independent claims. Further advantageous arrangements and embodiments of the invention are set forth in the respective subclaims. Reference should now be made to the appended claims.

In short words, the inventional method tracks the paths a user routes through the navigation topology in order to learn from what the user does. During future sessions when navigating at the same web portal's same node it blends-in recommendations that allow the user to go short-cuts to reach quicker the final destination page, or in graph terminology destination nodes, he is probably interested in, based on what the system has learned, when implementing the inventional method.

In a preferred embodiment a functional component tracks the user behavior, a further component analyzes the user behavior, and a third component displays the recommendations in the web browser user interface and provides the redirection functionality associated with each recommendation (short-cut).

The inventional method provides a set of useful short-cuts for each page in the navigational structure, which short-cuts try to reduce the time and effort needed to navigate through the usual paths as it is known from prior art. The sets of short-cuts are generated dynamically during user navigation, wherein the short-cuts are specifically generated dependant on the current node and preferably also dependant on the past visited

- 5 -

nodes. One ore a list of short-cuts is then displayed in a prominent position, preferably in the portal theme.

In order to do that the inventional method builds a model which delivers information about the number of times each link in the navigational structure was followed by the user. This information is used to calculate the probability that the user navigates to a particular page P conditioned by the fact that he is currently at another page Q. Then it predicts which may be the most probable ultimate destination in the current navigation of the user. In this way, the inventional recommendation method tracks, which are the most followed paths for each user, and places recommendations depending on where the user currently navigates, in order to reduce the number of clicks needed to navigate through the most popular navigation paths.

There is a trade-off between the usefulness of the inventional method and the number of short-cuts recommended. The recommendations provided by the inventional method are tailored to a specially focus a prediction of target pages that are quite far away in navigational distance, but which still have a high probability to be the final destination. Thus, basically the inventional method tries to achieve the highest expectable saving in number of clicks for traveling through tree-like navigation trees. When the recommendations provided by the inventional method are a list of short-cuts blended into the theme of the currently used web page, the short-cuts can be reached very easily by the user. Then only a simple click on a short-cut navigates the user to its desired final destination page, or portlet at this page.

According to its broadest aspect the present invention

- 6 -

discloses a method and system in a web portal environment for assisting a user during navigation in a given hierarchical navigation topology comprising nodes and edges connecting between said nodes, in which topology a web portal page corresponds to a predetermined topology level and displays one or more portlets, which comprises the steps of:

- a) during navigation of a user and for each page comprised of the navigation in the topology of the web portal automatically sensing information about the navigational behavior of the user,
- b) deriving page hits and page target hits from the sensed information, wherein page hits are defined for pages clicked for reaching another page, and page target hits are defined for pages a user clicks in order to perform a page-specific task or to receive some page-specific information,
- c) using a predefined metric based on the sensed information for generating a utility ranking for the web pages comprised of the user navigation, wherein the utility ranking calculation of a page includes a time-related weighting prioritizing newly used pages,
- d) tracking the current navigational position during the user's navigation,
- e) calculating for the current position and a plurality of target pages a probability measure to navigate to a respective one of the target pages, wherein the target pages are obtained by the utility ranking,
- f) calculating an expected saving factor for each target page based on the target page utility and on the navigational distance between the current page and the target page,
- g) defining at least one shortcut including the respective direct link to at least one of the target pages having a high expected saving factor,
- h) displaying the short cut as a clickable link to the user.

- 7 -

To sense information is meant to include to use all particular interfaces to "sensors" and read the information provided by them. Thus, time information is sensed from the system clock, date information is sensed from any date source of the Web server system, the user Id is sensed from a preceding log-in procedure, etc.

Examples for metrics in use are:

The amount of times (i.e. how often) a user has interacted with an element of a page, or
the amount of times (i.e. how often) a user was viewing a page, or
the frequency of visits of a page, or
the current location of a user, or
the hit time, month and day of week, etc..

The term "utility" is meant to denote the importance of a web page for a user. This importance is tried to be captured by above metrics. Of course, a further, importance parameter can also be used which is independent of the above metrics, if that seems useful to any use case in question.

The time-related weighting reflects the use changes which influence the user behavior slightly but usually steadily.

The current navigational position may be defined as a page URL, or a portlet URL.

A probability measure may be defined mathematically as a relation between specific page hits and overall page hits. Further, it may also be defined in a different way including some extensions to this approach. It follows the basic idea to

- 8 -

reflect the importance of a web page.

An expected saving factor reflects the user's work to reach a desired target page or target portlet and its relevance for the user, as long as this may be determined by means of a computer program. Certain aspects may be stressed in relation to others, for example, those portlets which are selected by a user probably in time-critical situations.

3. BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example and is not limited by the shape of the figures of the drawings in which:

Figure 1 is a schematic diagram illustrating the structural components of a prior art portal server including lines of interactions;

Figure 2 is a schematic block diagram representation of a portal server according to a specific embodiment of the invention illustrating structural components implemented therein;

Figure 3 is a schematic diagram illustrating a zoom-in into box 175 of Figure 2,

Figure 4 is a schematic interaction diagram illustrating a sequence of actions performed by a user (actor) and the components implemented by the inventional method, as illustrated in Figure 3,

Figure 5 is a schematic diagram illustrating the control flow

- 9 -

of a preferred embodiment of the inventive method during observation of the user behavior when navigating on a web portal topology.

Figure 6 is a schematic control flow diagram illustrating the essential steps performed in a preferred embodiment of the inventive method, when short-cuts are generated and displayed to the user dynamically during navigation, and

Figure 7A to 7H are formulae, describing a preferred implementation of the calculation of expected savings according to the invention.

4. DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

With general reference to the figures and with special reference now to figure 2 a preferred embodiment of the present invention will be described in more detail next below.

Figure 2 shows an additional inventive component 175, built-in into a prior art portal server and cooperating closely with the aggregation component 170 via a respective programming interface.

With particular reference to Figure 3 the before-mentioned aggregation component 170 controls the further functionalities depicted in boxes 172, 174, 176, 177 and 178. It further displays a given topology of a portal to a portal-user.

A web-usage mining component 176 has a read interface to the web-server's inbox, which receives a user request resulting from a user clicking somewhere at a link visible on the currently used web-page. Further, the component 176 has an interface to the web-server's clock in order to store the

- 10 -

arrival times of a user request. Via these interfaces the web-usage mining component is notified on each user interaction performed by a user when navigating certain portal topology. Component 176 calculates the before mentioned hit-values and target hit-values, associated with a given web-page and associated with a given context and invokes the database storage component 174 in order to perform a complete set-off sensing information, while the user is interacting with the web-portal.

A context observer component 177 is provided according to this embodiment in order to determine context attributes for defining the currently used context. In order to do that component 177 has respective functional interfaces in order to be unable to request exact time information, date information, to identify the device type currently in use by the web-portal-visitor, and possibly further information derivable or input by any other sensor device, like for example a GPS sensor element. Component 177 feeds any context information to the web-usage mining component 176 in order to enable component 176 to store current sensing information with the respective correct context.

A database storage component 174 is provided according to this preferred implementation for each user which stores the difference between a user-specific topology and the original topology which is also basically stored on the web server physically. Preferably, for this task an additional database table of a prior art portal database is provided. This table is denoted as 178, where the database is denoted as 128 in Figure 1 and 2. Each time, when a user has successfully logged-in the user-specific data stored in this specific

- 11 -

database table 178 is read and evaluated by a dynamic assembly transformation component 172 in order to reconstruct the user-specific navigation topology.

With further reference now to Figure 4 the interaction between the portal-visitor (user or actor 410) and the before mentioned components 170 to 177 are described in more detail next below:

The user 410 interacts with the web-portal web-pages. This interactions includes multiple repeated actions, for example clicking a link, clicking a portal, clicking any resource which can be clicked on a web-page. In order to illustrate the plurality of interactions, the respective arrow 420 is symbolized as a loop symbol. For each interaction the following procedure will be performed:

A current interaction is notified to the web-usage mining component 176. Such notification includes a user-ID, which is obtainable after a log-in of the current user and it includes information which node of which page was clicked and when this has occurred. Thus, a respective data set includes user-ID, node-ID, page-ID and time and date information. Optionally, of course it may include also device type information from which one can derive if the user uses PDA or a notebook or a desktop computer, respectively. This is done in step 430.

All data is send to the context observer component 177, see step 440, in order to determine the actually used context, see step 440 and above step 430. When the data collection is complete, it is stored in a step 450 within the database 178, by the database storage component 174.

- 12 -

Then, in a next step the same user is assumed to log-in again, see step 460. In this case the aggregation component 170 invokes a mechanism which evaluates the user-ID, if available the currently used device-ID and other context information (see above).

Then, aggregation component 170 requests to embed recommendation (to be generated according to the invention) into the theme of the current web page, see step 465.

Transformation component 172 will then request, step 470, the current determination of the currently used context. It directs a respective request to the context observer component 177, which reads the current context data from the last request header, and from other input sources, such as system time, in order to deliver time stamps, etc.. It receives this context on a subsequent step 475.

The recommendation component 172 then reads the stored and filtered interaction sequence from the database, see step 480, 485, and generates the recommendations. Details, how these recommendations are generated, are described in more detail further below.

Then the recommendation component returns those recommendations to the aggregation component 170 which displays it at the current web page theme.

Next and with reference to Figure 5, details of the user behavior observation are described:

In a step of sequence of steps 510 a user navigates to a node part of the navigation topology. This node part can be any

- 13 -

clickable item on the page. Then, a check 520 is performed testing if the user interacts on a certain, selected node, for example it is tested if he clicks a portlet. In the NO case a so-called retention time is calculated which is associated with this specific selected node based on the time associated with the click done in step 520 and the time of the next, subsequent click associated with a user request and received at the inbox of the web server. If the retention time which will be normally the difference of both times is higher than a pre-calculated median retention time it is branched to step 530. If not it is branched back to step 510. Also in the YES case of decision 520 step 530 is entered. In step 530 the target hit value for the selected node is calculated newly by using the web-usage mining component 176.

Then in a next step 535 the currently context attributes are determined using the context observer component 177. Finally, in a step 540 the new target hit value is stored in the proper context partition using the database storage component 174. Then it is branched back to step 510 in order to observe the next user interaction.

Figure 6 is a control flow diagram illustrating the most essential steps of the inventional method when generating a short-cut list for a given navigation node. With reference to Figure 6 the control-flow when creating short-cuts to target pages will be described in more detail:

Aggregation component 170 is assumed to traverse a given portal navigation topology. This is done in close relationship to the user actions done and sent by the inventional method, see step 610.

For each page the following actions are performed:

- 14 -

In step 620 a loop condition is tested in order to do the following loop body for all web pages the user navigates to:

In a step 630 the aggregation component 170 invokes the recommendation component 172.

Then in a next step 640 the user-ID of the currently logged-in user is detected and evaluated, in step 650 the respective current context attributes (date, time used client device) is determined by using the context observer component 177.

Then in a next step 660 the stored target hit values for all nodes for the current user and the current context partition are read and evaluated.

In step 670 a function calculate "expected saving" is invoked. The output of this function is a list of target nodes of the topology. In a further step 675 this list of target nodes is transformed into a clickable link to a respective target node, thus, possibly a particular link information is provided. By that, a respective short-cut list is generated, wherein each link corresponds to a respective short-cut. Then, in a step 680 those short-cuts are displayed preferably in the theme of the current web page.

Next and with reference to figures 7A to 7H, details are given, describing a preferred implementation of the calculation of expected savings, as mentioned above.

First it is proposed to view the user's navigation through the Portal as a Markov chain. With this model, we assume that pages the user will visit in the future are determined by the

- 15 -

current page and the recent visiting history.

Thus, a n -order Markov model takes into account the current page and the last $n-1$ visited pages to determinate the probability of each page to be viewed next.

This model is trained by counting how many times a page p_i was visited, after having traversed a determinate path p_{i-1}, \dots, p_{i-n} . The probability of following a link p_{i-1} to p_i is then calculated as depicted in figure 7A.

Obviously, the complexity of this model and therefore the amount of resources it takes to make the required computation grow with the order of the model. It is suggested that using a first order Markov model provides good results, being even not worthy the benefits of using a higher order if one takes into account the increase in complexity of the model. Thus, in this embodiment, a first order Markov model has been used.

Until this point, the mechanisms explained to model the user were static with the time. That means that given a log file they produce a model of the user from the available data, which is only valid under the circumstances present when these logs were recorded. As the behaviour of the users change with time these models can become obsolete, implacating that some kind of recalculation with an actualised version of the log file would be necessary. Another possibility for a model to become obsolete is a restructuring of the content by the administrator. In that case, a recalculation of the model would be necessary every time a change in the content structure is performed.

To add time sensitivity, a time-weighting mechanism with the

- 16 -

following parameters has been designed:

T-relevancy: this is the time we consider the data from the log file to be valid. Thus, data older than that are not taken into account in the model.

N-slots is the number of slots in which the relevancy time is split. These slots are then used to weight the data from the log file. Thus, data recorded in recent time slots are given more weight than data from older ones.

This mechanism tries then to make an estimation of the hit rate for every page and link, instead of just counting the absolute number of hits. The hit rate is calculated by performing a linearly weighted sum of the hit rate estimations for each time slot. So, the hit rate for an element e in a time slot i can be calculated as depicted in figure 7B, , where

$T\text{-slot} = T\text{-relevancy} / N\text{-slots}$ is the time duration of a slot, wherein e is the element (page or link) for which we are trying to calculate the rate and $hits(e,i)$ is the number of hits this element has received in the time slot i .

From now on, we assume slot 0 is the most recent one and slot $N-1$ the oldest one.

Having performed the hit rate calculation for each time slot, we can get a global hit rate calculation by performing a linearly weighted sum of them as depicted in figure 7C:

As the weighting factor decreases with i , more recent hit rate calculations are given more weight than older ones. The constant quotient multiplying the sum works just a normalization factor.

- 17 -

By substituting equation of Figure 7B into equation of figure 7C we finally get a formule according to figure 7D.

As time goes by slots are shifted. Older slots are eliminated giving then space for new ones. The rate estimation is recalculated every time a new slot is added, being then the model always updated to the real situation.

The probability calculation in the previous static models is then performed using this rate estimation instead of the absolute value of hits per page / link.

In the case of the unconditional model a formula yields according to figure 7E.

In the case of the first order Markov model a formula yields according to figure 7F.

If we take a look at the equations of Figure 7D, 7E, 7F we can see that the constant quotient multiplying the Sum in figure 7D becomes irrelevant when performing the probability calculation. Thus, this parameter can be neglected to get an optimized version of the mechanism and just calculate a time-weighted version of the page or link hits as follows from figure 7G.

The time duration of a slot is a very important factor for the estimation of an element's hit rate. Choosing a very short interval of time brings to a really quick adaptation since the model is updated every T-slot, but the variance of the model increases. That means that the estimation's quality is reduced, since quick variations in the user behaviour have a huge repercussion in the model. In normal conditions, slots

- 18 -

should have a duration around one day.

Obviously, the time of relevancy is also a key factor here. It determines not only the period of time during which actions performed by the user are considered as still valid, but it also determines the

number of slots $N\text{-slots} = T\text{-relevancy} / T\text{-slot}$

for which the hit rate is calculated in order to get an estimation of the global rate. Thus, the higher the number of slots the lower is the variance of the model.

The selection of a value for the T-relevancy parameter depends on the type of Portal. So, very dynamic Portals where content or services are outdated quickly and replaced by new ones should have short relevancy times, while Portals with a very static structure should have longer ones.

With respect now to the click distance d between a current, user-navigated node and a target node in a number of clicks only, without including mouse pointer movements, to which target node a short-cut is to be generated by the inventional method, it should be taken into consideration that, the further away a potential target node is from the front page, the more difficult is it to reach that target node. In order to integrate this aspect into the calculation of expected savings, a formula to calculate the expected savings is provided by figure 7H, wherein the links saved by the shortcut equals to: $\text{distance} - 1$, and wherein p is the probability a user navigates from the current page to the target node.

In this formula, the distance is ``discriminated'' in front of the probability: while the probability of reaching a node one level deeper in the structure decreases exponentially, the

- 19 -

distance between this node and the current one increases only linearly, in case of exponent $n=1$. In this way, nodes near to the current page are more likely to be recommended than others far away from it.

On the other hand, by taking $n > 1$ we start giving more importance to the number of links saved by the shortcut. That has as a consequence that the system makes more aggressive recommendations, that is, shortcuts to nodes which are further away from the current page.

The skilled reader will appreciate that the before-mentioned embodiments showing certain procedures and proposed sequences of steps are primarily given for demonstrating the inventional method by way of example. Thus, it will be understood that those examples may be varied in various aspects like for example the ordering of steps, the degree of parallelization between steps, the implementation details of the single steps, the form in which data is stored, the way how the grouping of portlets is achieved, etc., and respective further embodiments can be obtained.

The present invention can be realized in hardware, software, or a combination of hardware and software. A Portlet arrangement tool according to the present invention can be realized in a centralized fashion in one computer system, or in a distributed fashion where different elements are spread across several interconnected computer systems. Any kind of computer system or other apparatus adapted for carrying out the methods described herein is suited. A typical combination of hardware and software could be a general purpose computer system with a computer program that, when being loaded and executed, controls the computer system such that it carries

- 20 -

out the methods described herein.

The present invention can also be embedded in a computer program product, which comprises all the features enabling the implementation of the methods described herein, and which - when loaded in a computer system - is able to carry out these methods.

Computer program means or computer program in the present context mean any expression, in any language, code or notation, of a set of instructions intended to cause a system having an information processing capability to perform a particular function either directly or after either or both of the following

- a) conversion to another language, code or notation;
- b) reproduction in a different material form.

- 21 -

C L A I M S

1. A method in a web portal environment for assisting a user during navigation in a given hierarchical navigation topology comprising nodes and edges connecting between said nodes, in which topology a web portal page (125) corresponds to a predetermined topology level and displays one or more portlets (120), said method comprising the steps of:
 - a) during navigation of a user in the topology of said web portal automatically sensing (520, 525) information about the navigational behavior of said user for each page comprised of said navigation,
 - b) deriving (530) page hits and page target hits from said sensed information, wherein page hits are defined for pages clicked for reaching another page, and page target hits are defined for pages a user clicks in order to perform a page-specific task or to receive some page-specific information,
 - c) using a predefined metric based on said sensed information for generating a utility ranking for the web pages comprised of the user navigation, wherein the utility ranking calculation of a page includes a time-related weighting prioritizing newly used pages,
 - d) tracking (620) the current navigational position during said user's navigation,
 - e) calculating for said current position and a plurality of target pages a probability measure to navigate to a respective one of said target pages, wherein said target pages are obtained by said utility ranking,
 - f) calculating (670) an expected saving factor for each target page based on said target page utility and on the navigational distance between said current page and said

- 22 -

- target page,
- g) defining (675) at least one shortcut including the respective direct link to at least one of the target pages having a high expected saving factor,
- h) displaying (680) said short cut as a clickable link to said user.
2. The method according to claim 1, wherein a page hit is defined by a first click to a page and a subsequent click being done within a predetermined time range T-hit after said first click, and a page target hit being defined by a first click to a page and a subsequent click being done after a predetermined second time range T-target-hit having passed.
 3. The method according to claim 1, wherein a list of clickable shortcuts is displayed, which list is ordered according to the amount of the expected saving factor.
 4. The method according to claim 1, wherein said calculation of said expected saving factor includes the product of target page utility and navigational distance.
 5. The method according to the preceding claim, wherein said calculation of said expected saving factor includes a power of said product.
 6. The method according to claim 1, wherein said shortcut is displayed within the theme of the currently displayed web page.

- 23 -

7. The method according to claim 1, comprising the further steps of:
 - a) observing a context from the group of:
 - a1) current date,
 - a2) current time,
 - a3) currently used device type,
 - b) storing respective context information for a current user session,
 - c) using said context information as a filter for shortcut generation and/ or shortcut display.

8. A Web Application Server computer system in a web portal environment arranged for assisting a user during navigation in a given hierarchical navigation topology comprising nodes and edges connecting between said nodes, in which topology a web portal page (125) corresponds to a predetermined topology level and displays one or more portlets (120), characterized by comprising a functional component (175) implementing the steps of:
 - a) during navigation of a user in the topology of said web portal automatically sensing (520, 525) information about the navigational behavior of said user for each page comprised of said navigation,
 - b) deriving (530) page hits and page target hits from said sensed information, wherein page hits are defined for pages clicked for reaching another page, and page target hits are defined for pages a user clicks in order to perform a page-specific task or to receive some page-specific information,
 - c) using a predefined metric based on said sensed information for generating a utility ranking for the web pages comprised of the user navigation, wherein the

- 24 -

- utility ranking calculation of a page includes a time-related weighting prioritizing newly used pages,
- d) tracking (620) the current navigational position during said user's navigation,
 - e) calculating for said current position and a plurality of target pages a probability measure to navigate to a respective one of said target pages, wherein said target pages are obtained by said utility ranking,
 - f) calculating (670) an expected saving factor for each target page based on said target page utility and on the navigational distance between said current page and said target page,
 - g) defining (675) at least one shortcut including the respective direct link to at least one of the target pages having a high expected saving factor,
 - h) displaying (680) said short cut as a clickable link to said user.
9. A computer program for execution in a data processing system in a web portal environment arranged for assisting a user during navigation in a given hierarchical navigation topology comprising nodes and edges connecting between said nodes, in which topology a web portal page (125) corresponds to a predetermined topology level and displays one or more portlets (120), characterized by a functional component (180) implementing the steps of:
- a) during navigation of a user in the topology of said web portal automatically sensing (520, 525) information about the navigational behavior of said user for each page comprised of said navigation,
 - b) deriving (530) page hits and page target hits from said sensed information, wherein page hits are defined

- 25 -

for pages clicked for reaching another page, and page target hits are defined for pages a user clicks in order to perform a page-specific task or to receive some page-specific information,

c) using a predefined metric based on said sensed information for generating a utility ranking for the web pages comprised of the user navigation, wherein the utility ranking calculation of a page includes a time-related weighting prioritizing newly used pages,

d) tracking (620) the current navigational position during said user's navigation,

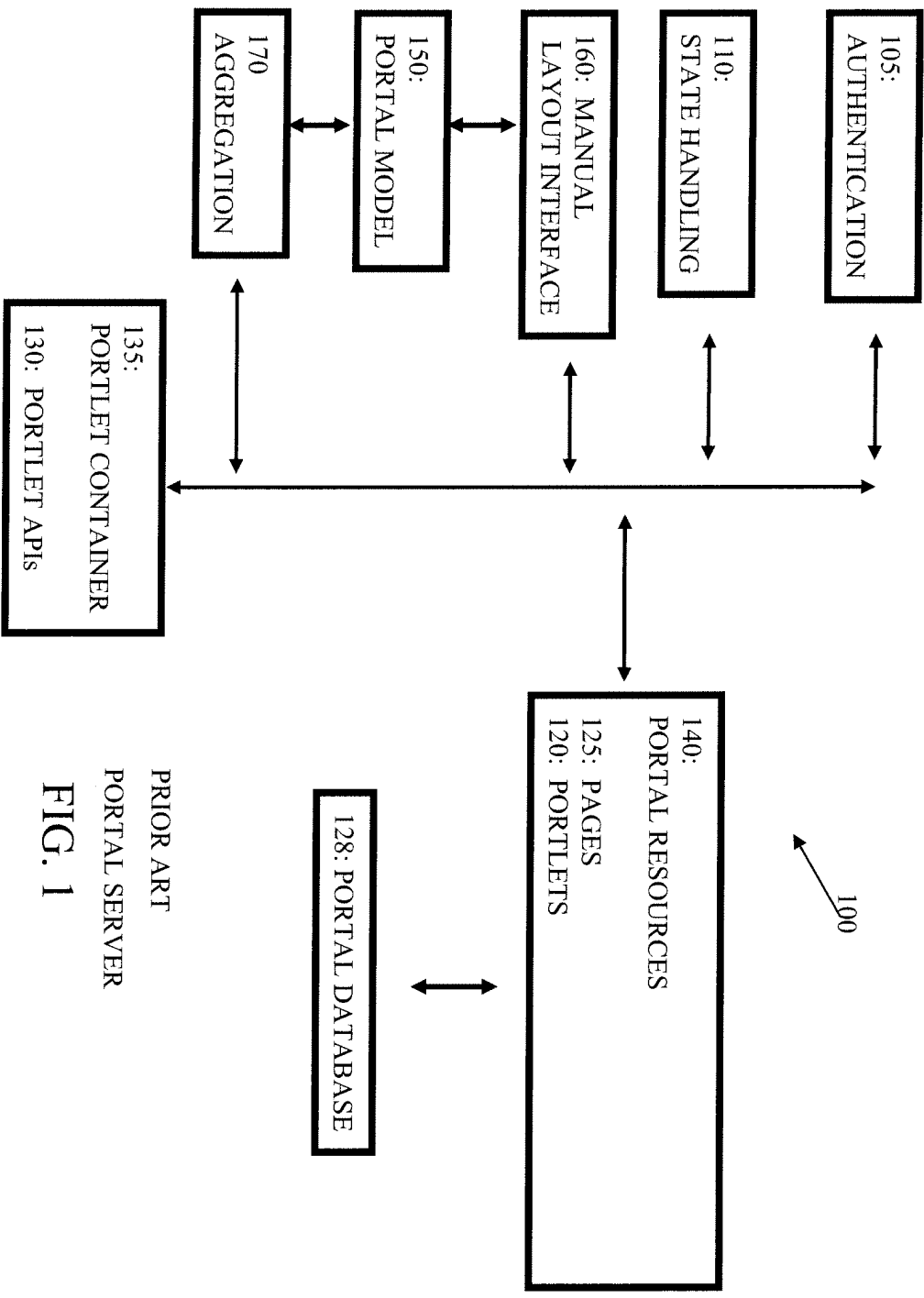
e) calculating for said current position and a plurality of target pages a probability measure to navigate to a respective one of said target pages, wherein said target pages are obtained by said utility ranking,

f) calculating (670) an expected saving factor for each target page based on said target page utility and on the navigational distance between said current page and said target page,

g) defining (675) at least one shortcut including the respective direct link to at least one of the target pages having a high expected saving factor,

h) displaying (680) said short cut as a clickable link to said user.

10. A computer program product stored on a computer usable medium comprising computer readable program means for causing a computer to perform the method of anyone of the claims 1 to 7, when said computer program product is executed on a computer.



PRIOR ART
 PORTAL SERVER
FIG. 1

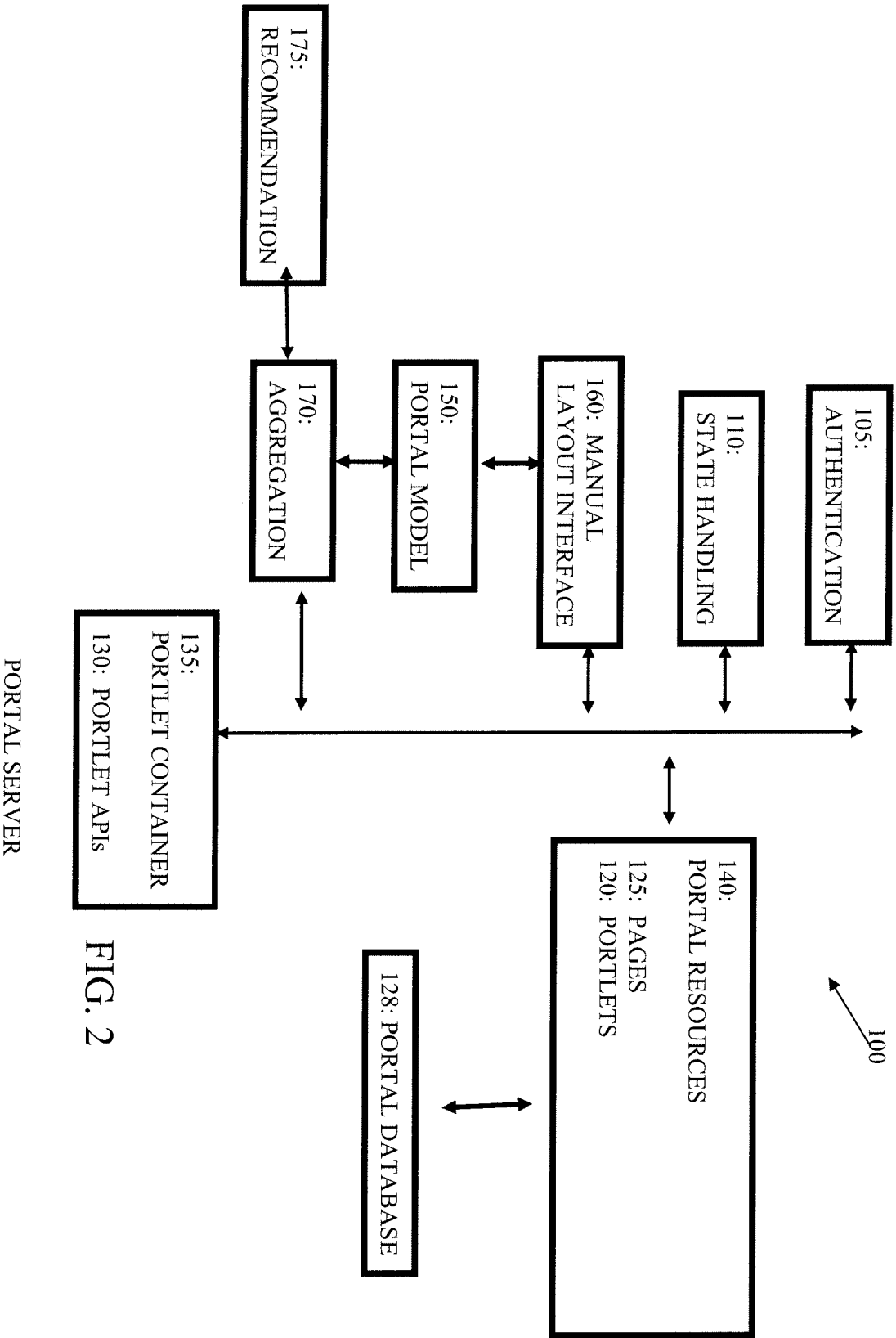


FIG. 2

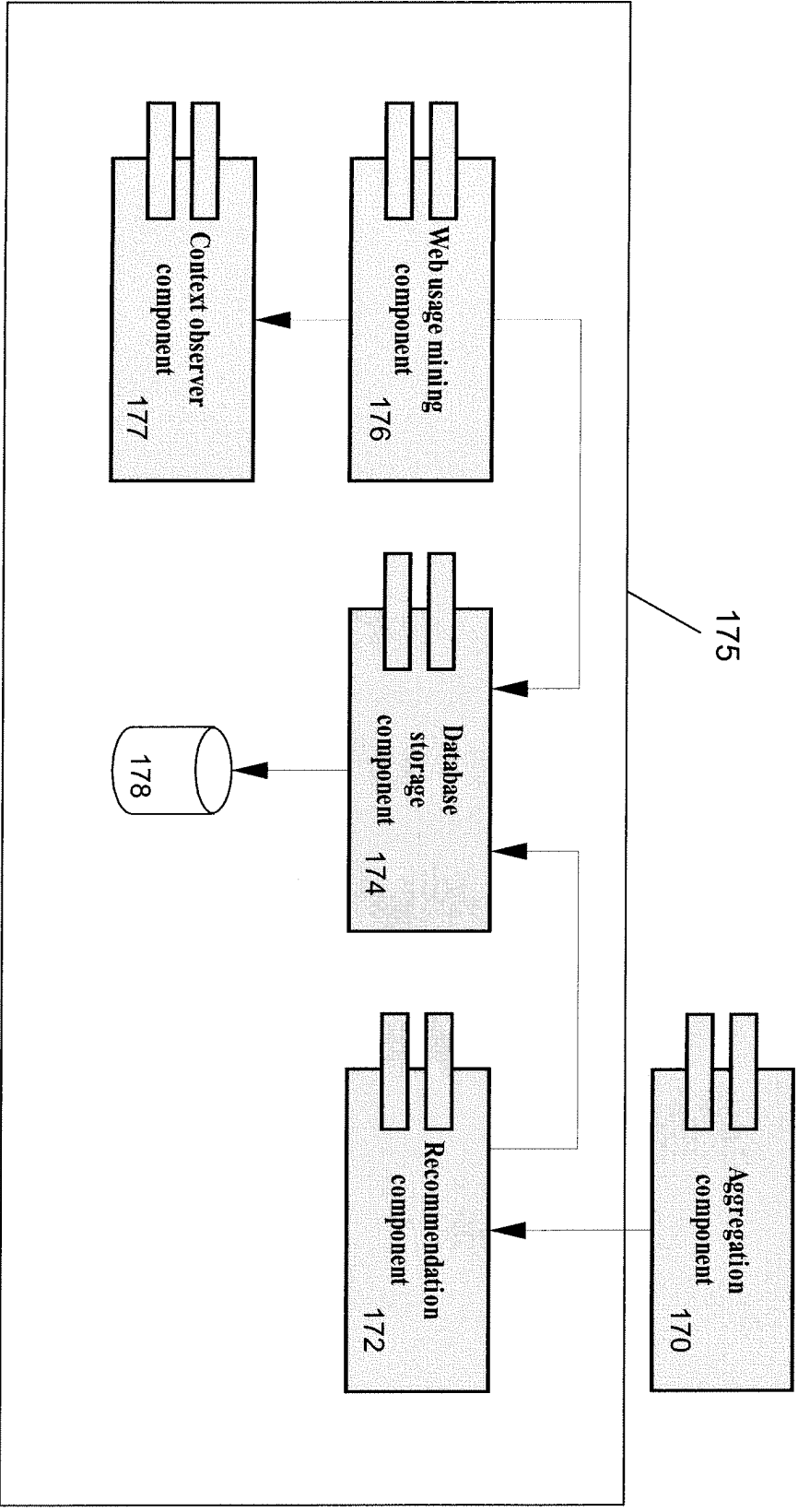
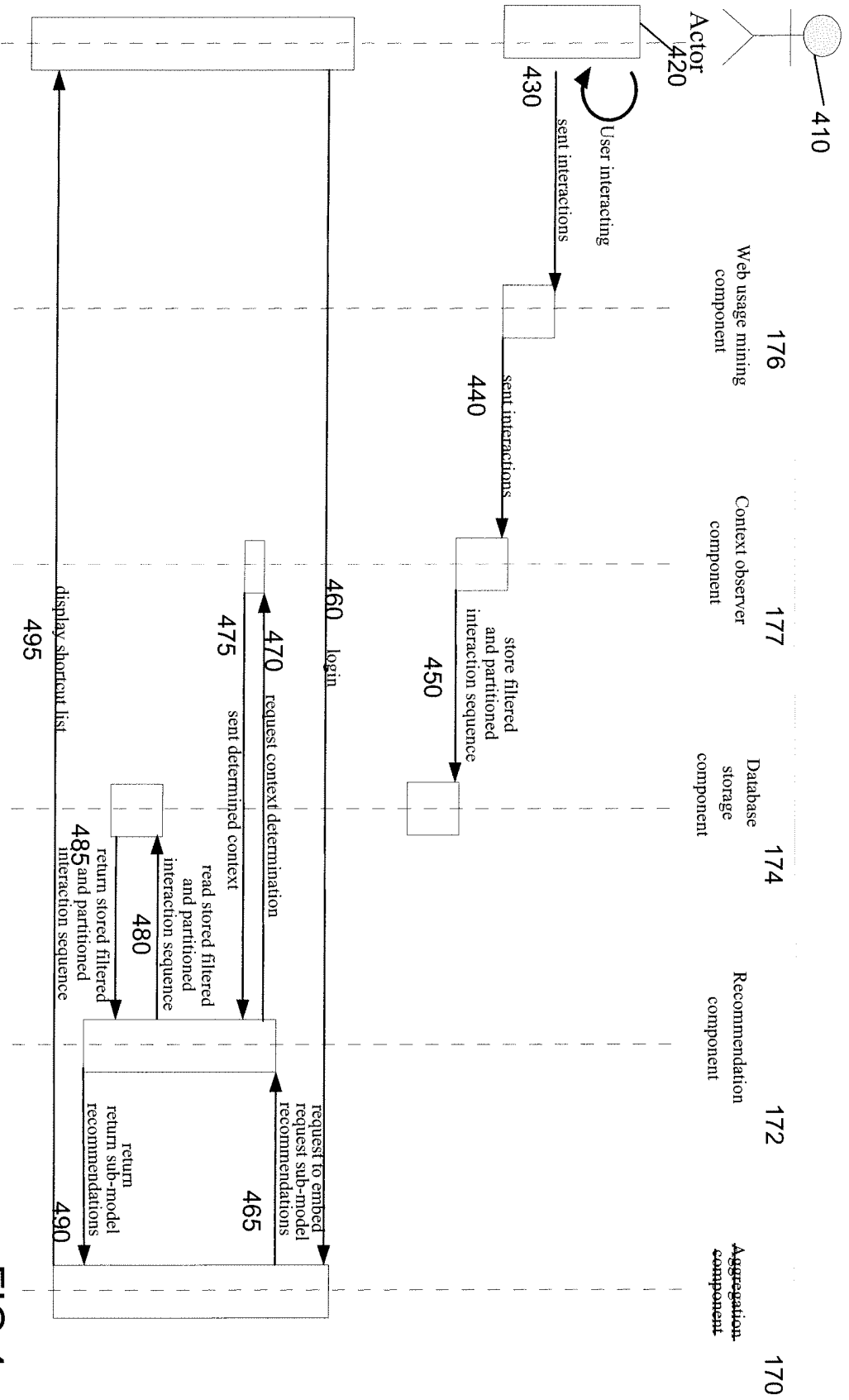


FIG. 3

FIG.4



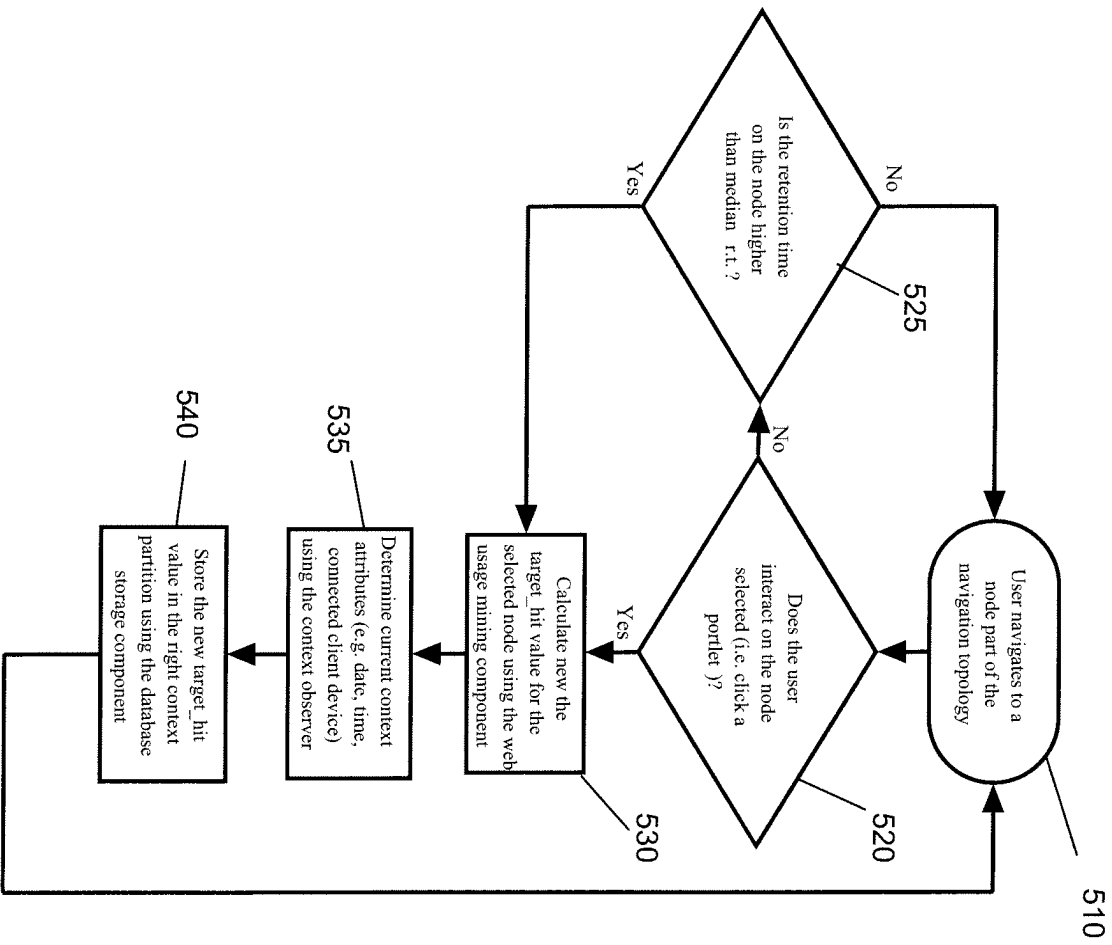


FIG. 5

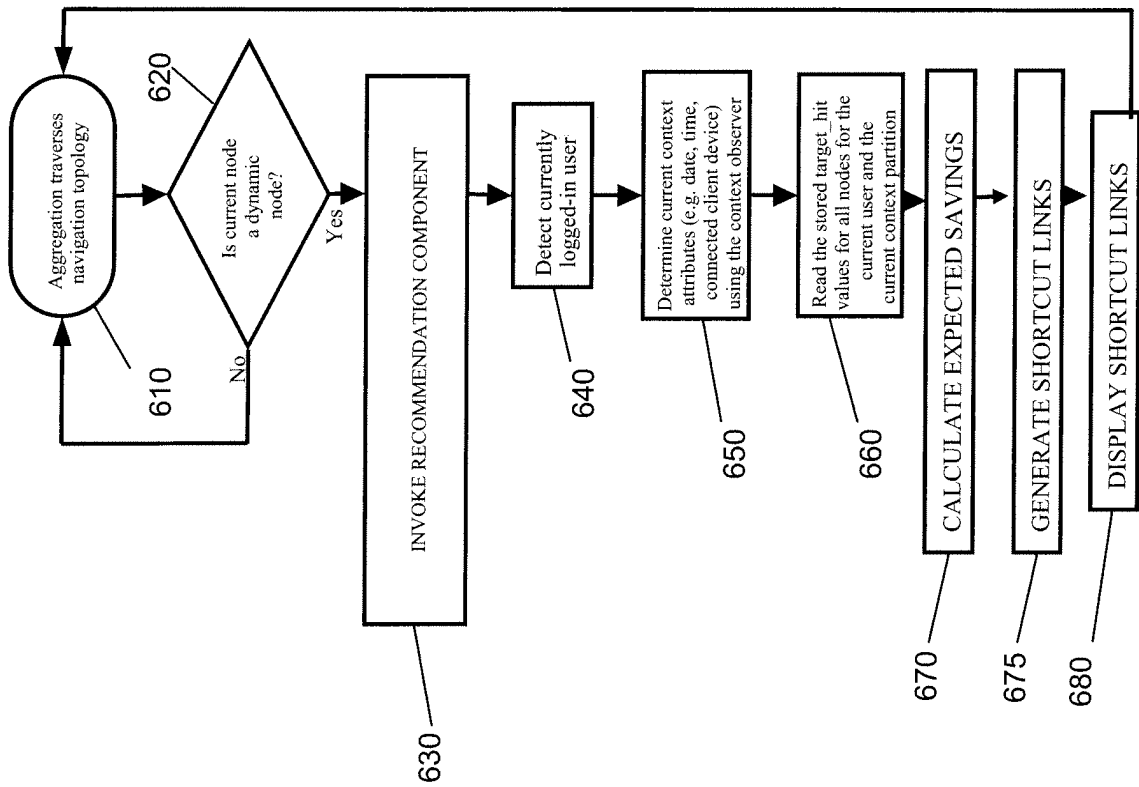


FIG. 6

$$\mathcal{P}(p_i | \langle p_{i-1}, \dots, p_{i-n} \rangle) = \frac{\text{number of times } p_i \text{ followed } \langle p_{i-1}, \dots, p_{i-n} \rangle}{\sum_j \text{number of times } p_j \text{ followed } \langle p_{i-1}, \dots, p_{i-n} \rangle} \quad \text{FIG.7A}$$

$$\hat{R}_{hit}(e, i) = \frac{hits(e, i)}{T_{slot}} \quad \text{FIG.7B}$$

$$\hat{R}_{hit}(e) = \frac{1}{(N_{slots} - (N_{slots} - 1)/2)} \sum_{i=0}^{N_{slots}-1} \frac{N_{slots} - i}{N_{slots}} \cdot \hat{R}_{hit}(e, i) \quad \text{FIG.7C}$$

$$\hat{R}_{hit}(e) = \frac{1}{T_{slot} \cdot (N_{slots} - (N_{slots} - 1)/2)} \sum_{i=0}^{N_{slots}-1} \frac{N_{slots} - i}{N_{slots}} \cdot hits(e, i) \quad \text{FIG.7D}$$

$$\mathcal{P}(p_{i+1} = q) = \frac{\hat{R}_{hit}(q)}{\sum_j \hat{R}_{hit}(p_j)} \quad \text{FIG.7E}$$

$$\mathcal{P}(p_i | p_{i-1}) = \frac{\hat{R}_{hit}(p_{i-1} \rightarrow p_i)}{\sum_j \hat{R}_{hit}(p_{i-1} \rightarrow p_j)} \quad \text{FIG.7F}$$

$$\text{weighted_hits}(e) = \sum_{i=0}^{N_{slots}-1} \frac{N_{slots} - i}{N_{slots}} \cdot hits(e, i) \quad \text{FIG.7G}$$

$$\text{expected_savings} = p \cdot (d - 1)^n \quad \text{FIG.7H}$$

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2007/056475

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04L29/08 G06F17/30

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

B. FIELDS SEARCHED

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

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/267869 A1 (HORVITZ ERIC [US] ET AL) 1 December 2005 (2005-12-01) paragraphs [0003] - [0011] paragraphs [0014], [0015] paragraphs [0039] - [0042] paragraphs [0049] - [0062] paragraph [0076] claims 1-7 figures 15-17	1-10
A	US 6 959 319 B1 (HUANG ANITA WAI-LING [US] ET AL) 25 October 2005 (2005-10-25) abstract figures 1,2,5A,5B,8 column 1, line 60 - column 2, last line column 3, line 40 - column 5, line 20 column 5, line 63 - column 6, line 52 ----- -/--	1-10

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

* & * document member of the same patent family

Date of the actual completion of the international search

14 November 2007

Date of mailing of the international search report

22/11/2007

Name and mailing address of the ISA/
European Patent Office, P.B. 5816 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Buhleier, Rainer

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2007/056475

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>WO 02/091154 A (CHANGINGWORLDS LTD [IE]; SMYTH BARRY JOSEPH [IE]; COTTER PAUL [IE]; MC) 14 November 2002 (2002-11-14) abstract page 1, line 10 - line 28 page 6, line 10 - page 8, line 10; figure 3</p>	1-10
A	<p>US 2005/177613 A1 (DRESDEN SCOTT [US]) 11 August 2005 (2005-08-11) abstract paragraphs [0001], [0002] paragraphs [0035], [0036] paragraphs [0044], [0047], [0048] figures 1,2,3A-C</p>	1-10
A	<p>WILL R ET AL: "WebSphere Portal: Unified user access to content, applications and services" IBM SYSTEMS JOURNAL, IBM CORP. ARMONK, NEW YORK, US, 26 April 2004 (2004-04-26), pages 420-429, XP002356355 ISSN: 0018-8670 abstract page 442 page 427, right-hand column, paragraph 3 - page 428, right-hand column, paragraph 1 figures 4,5</p>	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2007/056475

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005267869	A1	01-12-2005	US 7203909 B1 10-04-2007
			US 2005278323 A1 15-12-2005
			US 2006004763 A1 05-01-2006
			US 2006004705 A1 05-01-2006
			US 2005278326 A1 15-12-2005
US 6959319	B1	25-10-2005	NONE
WO 02091154	A	14-11-2002	CA 2445925 A1 14-11-2002
			CN 1714334 A 28-12-2005
			EP 1412844 A2 28-04-2004
			JP 2004531818 T 14-10-2004
			US 2003065706 A1 03-04-2003
			US 2003063129 A1 03-04-2003
US 2005177613	A1	11-08-2005	NONE