Personalized Information Retrieval Framework

Dongpyo Hong, Yun-Kyung Park, Jeongwon Lee, Vladimir Shin, and Woontack Woo

***Abstract*—In this paper, we propose a framework of personalized information retrieval system for a wearer (a person who is equipped with a wearable computer) in ubiquitous computing environments. In ubiquitous computing environments, personalized information retrieval is indispensable for a wearer because desired or undesired information will be flooded around the wearer. Although there have been many research activities on information retrieval in wearable computing and/or ubiquitous computing environments, personalization has not been recognized significantly. Accordingly, all the retrieved information is exposed to the wearer regardless of his/her situations or conditions. In this regard, the proposed framework enables a wearer to retrieve the personalized information from objects. In the proposed framework, we exploit user's context as a fundamental element in retrieving the personalized information. The proposed framework consists of two conceptual stages (object selection and information presentation) and each stage includes several components. In order to measure the effectiveness of the proposed framework, we introduce a measuring method and also realize a prototyping system for personalized information retrieval. Thus, we believe that the proposed personalized information retrieval framework is to leverage human-computer interactions for a wearer in ubiquitous computing environments.**

***Index Terms*—Context, Personalized Information Retrieval, Wearable Computing, Ubiquitous Computing**

I. Introduction

WITH the rapid progress of technologies in the areas of computers and communications, the future computing environments will support seamless interactions from ubiquitous computers and pervasive networking. That is, users could do just-in-time access to any (invisible) computer at any time and any where [1-3]. Consequently, this environment will require users to interact with computers through more natural and comfortable interfaces. Meanwhile, due to the explosion of the volume of available information for users to deal with in this computing environment, the information retrieval systems are indispensable to those who want to retrieve appropriate information with less effort in a given time.

There have been many research activities on information retrieval to find an appropriate document from database or libraries [4-6]. Over the past few years, however, it has been reported on the importance of information retrieval in wearable computing, mobile augmented reality (AR), and in ubiquitous computing environment. For instance, S. Julier, et al. (2000) proposed information filtering for mobile augmented reality because display information in AR is cluttered with too much information [7]. Insley (2003) figured out that one of difficult problems for AR is effectively managing large amounts of information [8]. Without a way to filter out garbage information, augmented reality would be utterly useless. Jones and Brown (2004) addressed that information retrieval may be facilitated by contributions from human-computer interaction studies and agent technology to determine how and when to deliver the information to the user or how best to act the user's behalf [9]. On the other hand, it is also an important feature to present the retrieved information appropriately to users afterward. For example, Sinclair, et al. (2003) suggested that Adaptive Hypermedia is a solution for dealing with complex, heavily structured information and the presentation of views of that structure to users [10]. Meanwhile, the relevant applications were mainly on the battle fields or tour guides. By way of example, S. Julier, et al. (2000) proposed that BARS (Battlefield Augmented Reality System) will ensure that only the most relevant information is displayed to the user at a particular time [11]. B. Bell, et al. (2003) proposed an AR system that could provide "information at a glance" to aid a mobile user exploring an unfamiliar environment [12]. The detailed reviews are shown in Section II. Whereas the previous research on information retrieval is mostly to find out appropriate documents from database regardless of user's context, the recent research takes into account user's context in the development of information retrieval systems. However, there are few frameworks to support personalized information retrieval system based on user's context.

In this paper, we propose a framework of personalized information retrieval system for a wearer (a person who is equipped with a wearable computer) in ubiquitous computing environments. Moreover, we suggest an effective means to exploit user's context through this framework. In the proposed framework, we separate the procedure of personalized information retrieval into two parts. The first part is object selection, which includes context transform and personalized object filter components. In the second part, personalized information filter and personalized information presentation

Manuscript received June 17, 2005. This work is supported by Seondo project of MIC, Korea. Personalized Information Retrieval Framework.

Dongyo Hong is with Gwangju Institute of Science and Technology (GIST), Gwangju, Korea (e-mail: dhong@gist.ac.kr).

Yun-Kyung Park is with Electronics and Telecommunications Research Institute (ETRI), Deajun, Korea (e-mail: parkyk@etri.re.kr)

Jeongwon Lee is with Department of Digital Contents, Sejong University, Seoul, Korea (e-mail: jwlee@sejong.ac.kr).

Vladimir Shin is with Gwangju Institute of Science and Technology (GIST), Gwangju, Korea (e-mail: vshin@gist.ac.kr).

Woontack Woo is with Gwangju Institute of Science and Technology (GIST), Gwangju, Korea (corresponding author, phone: +82-62-970-2226; fax: +82-62-970-2249; e-mail: wwoo@gist.ac.kr).

components are included. Besides these components, the proposed framework also includes context representation component. Before we introduce each component in the proposed framework, we assume that any information of an object as well as user's context, namely user's profile or preference, can be represented as a unified context. We name this process as **context representation** [13-17]. This basic assumption enables us easily to describe contextual information. Otherwise, contextual information is too complicated to utilize in personalized information retrieval system. In addition, the unified context representation allows us to access contextual information in a structured manner when we try to retrieve information from an object. Although we have the unified representation of contextual information, we need another transformation, **context transformation**, because the unified representations of context are not easily comparable. Thus, context transformation transforms the unified context to appropriate form such as Boolean or probability for easier comparison between user's context and information of an object. Owing to context representation and context transformation, we can compare information of an object with user's context in a consistent way. After context transformation, we are able to select the most relevant object in response to user's context, where the selection criterion is defined as **personalized object filter**. In the personalized object filter, we can exploit either deterministic or probabilistic selection criteria in order to filter out an object from various objects. After the selection of an object responding to user's context, we have another filtering to present only preferable information in the selected object through an appropriate output device. In fact, this procedure consists of **personalized information filter** and **personalized information presentation**. In this framework, however, we assumed that the output device of a wearer remains the same regardless of wearers. Lastly, we also introduce a measuring method for personalized information retrieval system, i.e. recall and precision measurement. As a matter of fact, recall and precision measurement is widely used in conventional information retrieval systems [4-6]. In this regard, we can project personalized information retrieval into a searching or classification problem, and thus can adopt various mathematical models like Boolean matching or Bayesian decision rule according to the applications.

Therefore, the proposed framework helps developers to implement personalized information retrieval system due to the structured components and easier exploitation of user's context. In addition, we believe that personalized information retrieval enables a wearer to do natural interactions with computing resources like augmented virtual object or information in ubiquitous computing environment. That is, the retrieved information is presented differently corresponding to user's context such as user's preference. Consequently, it leverages human-computer interactions of a wearer to retrieve personalized information from a large volume of information around the wearer. In order to realize the proposed framework, we exploit wear-UCAM as an empirical tool, a toolkit for context-aware applications, because context-awareness mechanism and other useful utilities are implemented for developers [18]. Furthermore, we deploy the implemented system for personalized information retrieval in ubiHome, a smart home environment test-bed, and experiment with it [2].

This paper is organized as follows. In Section II, we review the related works which address the necessity of information retrieval for wearable computing, mobile AR, and ubiquitous computing environment. In Section III, we explain the fundamental components for the proposed framework of personalized information retrieval with a specific example. In Section IV, we show the experimental results and effectiveness of the proposed framework through the introduced measuring method. Finally, we discuss future works in Section V.

II. BACKGROUND

In general, research on information retrieval has been focused on how to retrieve the requested information from a large volume of data, particularly documents, even images in the Internet (WWW), database, etc. Over the last decade, the concept of information retrieval has been moved to mobile computing, wearable computing, augmented reality (AR), and ubiquitous computing in order to retrieve only needed information because there will be too much information in such computing paradigms. In this section, we review the most relevant works about information retrieval, and then we extract fundamental components for personalized information retrieval system.

In [7], the authors proposed information filtering for mobile augmented reality because display information in AR is cluttered with too much information. Information filtering means to cull the information that can potentially be displayed by identifying and prioritizing the information that is relevant to a user at a given point in time. Thus, information can be classified based on the user's physical context as well as on their current tastes and objectives. In addition, the amount of information shown to a user about an object is inversely proportional to the distance of that object from the user. Although their information filtering framework includes well-defined components, the used functions in this framework are too general for developers to use them. Moreover, the objective properties and subjective properties in the state of user and object seem to be used as contextual information, but their representations are dependent on a specific domain such as battle field. Thus, the representations of user's context and object's context should be generalized so that the information filtering framework is applicable to other areas.

In [8], one of most difficult problems for AR is effectively managing large amounts of information. The cluttering junk information would make it difficult to find useful information and even become distracting to interface with real world tasks. Without a way to filter out garbage information, augmented reality would be utterly useless. In this regard, there are numerous methods that allow some level of filtration. One of the simplest techniques is to filter information based on location. However, local user's context like location information is not enough to filter out appropriate information from useless information. Thus, a set of user's context should be applied into information retrieval. Meanwhile, the most basic type of content is labels. They can be used to display simple yet important pieces of information. However, if their numbers become excessively large, the labels will become a problem, despite their individual simplicity. It is therefore necessary to have a method of automatic filtering labels to keep them under control, although it need not be terribly draconian.

In [9], the author addressed that information retrieval may be facilitated by contributions from human-computer interaction studies and agent technology to determine how and when to deliver the information to the user or how best to act on the user's behalf. To improve retrieval effectiveness in the future computing environments in terms of retrieval accuracy and user satisfaction, the key issue is the integration of technologies from information retrieval and context-awareness. However, current search engines take no account of the individual user and their personal interests or their physical context. Although the author did not mention explicitly the term 'wearable computer', we believe that personal and context information is potentially available for the retrieval process by a wearable computer. Meanwhile, the information should be delivered in a timely fashion since the information is based on the user's context and the context may change. Thus, it will be often useless to deliver information about a situation the user has just left. In this regard, information retrieval can be extended to incorporate recommendations based on preferences or profiles derived from the user where long-term information needs because the interests of the user are modified gradually over time as conditions, goals and knowledge change. Lastly, the personalization for retrieval is how the interests of the user are captured. The personalization system monitors the user's behavior and learns profiles from this.

In [10], Adaptive Hypermedia (AH) is a solution for dealing with complex, heavily structured information and the presentation of views of that structure to users. The AH system may maintain a user profile for each of the people interacting with it, initially based on user preferences. And the system modifies the content that they see and the paths that they take through the content accordingly updating the profiles as the users move through the information space. Users view information that is commonly presented to all users regardless of the context of their interaction, or any profiling, as non-adaptive. When the information to be displayed is adaptive, meaning that it changes according to the context of the viewing, most AR researchers use a dynamic approach. Thus, information retrieval should exploit contextual information, especially user's context. However, there is no relevant explanation about a way of representing user's context and exploiting it in the actual interactions.

In [11], the filtering is a logical extension and refinement of the aura used with the object distribution in order not to show all irrelevant information to users. To help the decision process and to take into account the variety of situations that can be encountered, the user can select a filtering mode according to his current mission. In addition, the paradigm of an information filter is presented by a decision mechanism that uses the user's location, the user's current goal and the properties of objects within the environment to deduce what information should be displayed. The filtering performed by BARS (Battlefield Augmented Reality System) will ensure that only the most relevant information is displayed to the user at a particular time. From this, we can project personalized information retrieval to a decision problem, and thus we can employ various mathematical approaches.

In [12], the authors presented examples from prototype augmented reality systems that show some of the ways in which information might be displayed, emphasizing the automated layout of overlaid graphics. In this approach, they focused on view management in accordance with user's context, especially user's location indoor and outdoor. However, they assumed that information filtering has been done and that everything being displayed should be displayed. Thus, we believe that information presentation is another important feature in personalized information retrieval system.

In this review, we observed that many researchers have addressed the importance of information retrieval in the future computing environments. Although most researchers point out the usefulness of user's context in information retrieval, there are few works to reveal how to exploit user's context and its detailed manipulation for personalized information retrieval system in such computing environments. As we have reviewed, there are many practical applications which show the effectiveness of information retrieval. However, there is little tendency to show the detailed explanations of the information retrieval algorithms supported by mathematical derivations and evaluations. These have inspired us to investigate on the framework for personalized information retrieval system with mathematical models.

III. Personalized Information Retrieval Framework

A. Overview

From the previous sections, we could define that personalized information retrieval is to retrieve appropriate information from a large volume of data (or information) with a wearer's context, namely preference or profile, and also to present the retrieved information appropriately based on the wearer's context in ubiquitous computing environment where any information could be augmented by anyone. As shown in Figure 1, information or property of any object and wearer can be represented as a unified context [13-17]. In addition, information retrieval should occur only if a wearer has a special interest on a specific object. Thus, the most relevant information to the wearer is retrieved from the selected object. However, the preferred information in response to the wearers' context within the retrieved information should be exposed to them. For example, let Wearer A and Wearer B are interested in the same object in the given environment. However, their preferences are supposed to be different. In conventional information retrieval system, the retrieved result of Wearer A is the same as the result of Wearer B because they are interested in the same object, that is, Query A is equal to Query B. In the proposed framework, however, the retrieved results of Wearer A and Wearer B are different because the queries are based on their context, i.e. Query A and Query B are differently treated. Therefore, personalized information retrieval is achieved in the proposed framework.

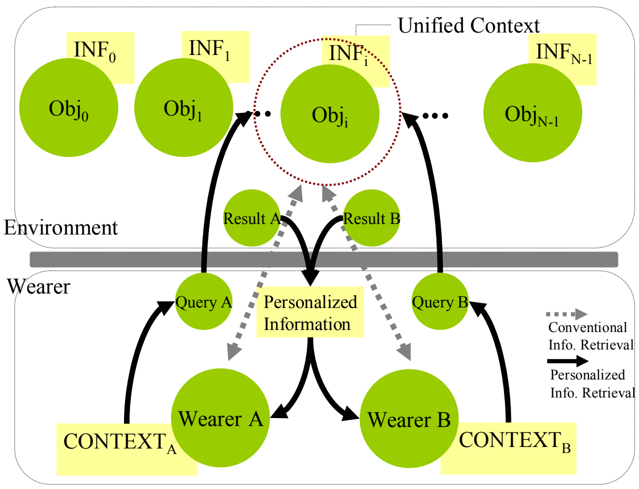
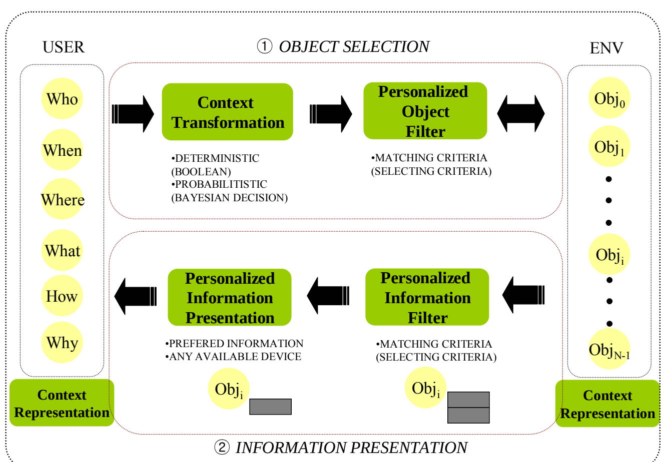


Fig. 1. Conceptual Diagram of Personalized Information Retrieval for Wearers in Ubiquitous Computing Environment.

Figure 2 shows the procedure of the proposed personalized information retrieval framework. We will explain the procedures and components of this framework in the following subsections.



Personalized Information Retrieval Framework

Fig. 2. Personalized Information Retrieval Framework.

B. Context Representation

There are many definitions of context among researchers [13-17]. However, we exploit the unified context as the form of 5W1H (Who, When, Where, What, How, and Why) to represent user's context and object's information (properties) [15]. Due to this unified representation of user's context and object's information as 5W1H, we are easily able to describe contextual information. Otherwise, contextual information is too complicated to utilize in the information retrieval system. In addition, the unified context representation allows us to access contextual information in a structured manner when we try to compare contexts, select an object, and retrieve information of the object in the proposed framework. Generally, the unified contexts can be implemented by key-value models, markup scheme models, graphical models, object oriented models, logic based models, or ontology based models [17]. Regarding the utilization of the unified context in the proposed framework, we internally employ temporary data structures to manipulate user's context and object's information effectively. In the framework, two temporary data structures are introduced; one is for user's context (**contextual query**), and the other one is for object's information (**object database**). Through the introduction of temporary data structures, we can simplify personalized information retrieval problem as a query from a database. In Section IV, we will show an example to help understand context representation. Figure 3 shows the specific implementation of contextual query from user's context and object database from available objects in the given environment.

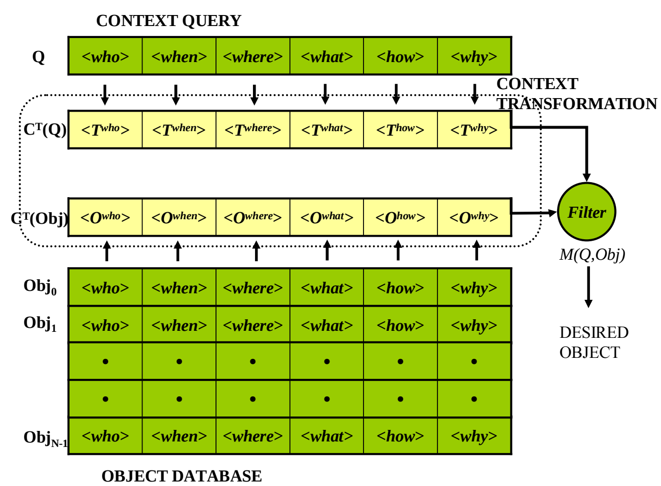


Fig. 3. An Example of Temporary Data Structures.

As shown in Figure 3, a user's context can be represented as **Q** and available objects can be represented as a vector of **Obj** respectively. Furthermore, their elements in the set are vector types. However, this internal data structure is nothing but data container of the user's context and of available objects, thereby we need to transform it into another form, which is explained in the next subsection.

C. Context Transformation

In general, users would always retrieve exact and perfect information to their requests from a given system. However, this is only possible when the request is exactly coincident and information in the system can be definitely identified as responding to the query. For example, the most obvious exact matching arises with a numerical representation of information. Although the context representation of 5W1H format has many advantages, it is hard to be used directly in the proposed framework because the context representation is just a data container. Therefore, user's context and object's information needs to be transformed to an appropriate format for easier comparison such as Boolean or numerical format without loss of advantages in the context representation. We define this transformation as **context transformation,** and it can be represented as follows.

$$X' = C^T(X) \tag{1}$$

where X is an arbitrary context, X' is the transformed context, and C^T represents context transformation.

In context transformation, any transformation method is applicable if the transformed contexts are definitely comparable to each other. In the proposed framework, either deterministic transformation or probabilistic transformation is possible. However, we exemplify Boolean transformation method in a deterministic transformation for the sake of simplicity. Typically the connectives in Boolean permitted are AND, OR, and NOT. In the Boolean model of retrieval, the similarity between an object and a query is based on the presence of terms in both the query and the object. In spite of the simplicity of Boolean model, they present a number of significant problems [4-6]. For instance, there is no good way to weight terms for significance. Regarding weight, the vector model based on *0-1*vector can be exploited, where each component is either *0* if the corresponding term is absent or *1* if the corresponding term is present in the object or query being represented. Whereas Boolean models have the disadvantage of not incorporating term weights, vector models have the disadvantage of not being capable of expressing logical connectives easily. Various attempts have been made to mediate Boolean and vector queries by the introduction of weighted or extended Boolean queries [19-21].

Thus, we can represent user's context as a weighted Boolean query as follows.

$$Q' = T_{w_1}^{who} * T_{w_2}^{when} * T_{w_3}^{where} * T_{w_4}^{what} * T_{w_5}^{how} * T_{w_6}^{why} \tag{2}$$

where $\boldsymbol{T}^{5W1H}$ are query terms, $w_{1\ldots6}$ are the weights (either 0 or 1), and * represents a Boolean or logical operation. Although sum of weight is equal to 1, the term of weight used in context transformation only indicate whether the transformed query includes terms or not. Thus, the total sum of weights need not be 1. As shown in (2), user's context can be transformed weighted Boolean query format, where each term of query is correspondent to each term of 5W1H. In case of context transformation of objects, however, we just take weight of each term as 1.

In general, the weighted Boolean operations are relatively complicated [5,19-21]. However, we can avoid the complicated weighted Boolean operations because each term of 5W1H is independent value. This is one of the advantages of the unified context representation in 5W1H format. Regarding context transformation, we will explain the way of using context transformation through an example in Section IV.

D. *Personalized Object/Information Filter*

Ultimately, the goal of information retrieval is to find a match between a given query and those objects that the user would like to retrieve in response to the query. The matching process is complicated by the fact that the query and the objects may have quite different forms. Through the context transformation, however, we are able to cope with this complication. On the other hand, a term in an object has been matched to a term in a query does not automatically mean that the object should be retrieved in response to the query. In the first place, the query probably contains more than one term, so the success or failure in matching each of the query terms must be considered. In the second place, an object which contains a given query term does not mean that the object is strongly related to the term. Therefore, matching criteria is a key factor in personalized information retrieval. After we briefly examine the types of matches that can be used, we will show a criterion as an example of **personalized object/information filter**.

In general, the matching criteria in information retrieval are categorized as follows: exact match, range match and approximate match [4-6]. The most obvious exact match situations arise with numerical information. Any information whose term values all match the values specified in the query will be retrieved; all other information will be rejected. An immediate extension of the exact match is the range match. In this type of matching, a range of acceptable values is given for each term. Range matches are possible on terms that have a natural order, for example numerical or alphabetical, so that it is meaningful to specify that a term be greater than some minimum value or less than some maximum value. An approximate match requires some way to measure how well a given object matches the query.

We focus on evaluating a match between an object (**Obj'**) and a query (**Q'**). In addition, we exploit the cosine measurement which is widely used as a matching algorithm in information retrieval systems [6]. In this measurement, we assumed that the object and query are represented as numerical vectors in *t*-space, that is **Q'** = ($q_0, q_1, \ldots, q_{t-1}$) and **Obj'** = ($obj_0, obj_1, \ldots, obj_{t-1}$) where q_i and obj_i are numerical weights associated with the keyword *i*. The cosine correlation is now simply defined as follows.

$$M(Q', Obj'_k) = \frac{\sum_{i=0}^{t-1} q_i obj_i}{\left(\sum_{i=0}^{t-1} (q_i)^2 \sum_{i=0}^{t-1} (obj_i)^2 \right)^{1/2}} \tag{3}$$

With the matching function (3), we can retrieve the object of interest as the following criterion. Suppose there are *N* objects in the environment, then the retrieving proceeds by calculating values *M(***Q'**,**Obj'**$_{\mathbf{k}}$*)*, where *k* is an arbitrary number in *N*. Through this calculation, the set of objects to be retrieved is determined. In mathematical point of view, this function is the inner product of the query and object vectors, where 1 represents the highest similarity. Thus, the function *M* transforms the angular measure into a measure ranging from 1 for the highest similarity to 0 for the lowest. In general, there are two approaches to retrieve objects of interest: threshold and rank position approach. However, we adopt threshold approach because it is intuitive and widely used. When the matching function is given a suitable threshold, we can retrieve the objects above the threshold and discard the ones below. If *Th* is the threshold value, then the retrieved object set is represented as $\hat{O} = \{\mathbf{Obj'_k} \mid M(\mathbf{Q'}, \mathbf{Obj'_k}) > Th\}$.

On the other hand, we have to keep the space or length, *t*, of the query and object as the same space in order to exploit the vector model. In the case of personalized object filter, this condition is kept because the user's context and object's context are represented in 5W1H formats, i.e. t is equal to 6. However, it is not natural to restrict the vector space of each term in the user's context and object's context as the same space. Thus, we simply count the number of query terms which an object contains when we filter out the preferred contents from the retrieved information in response to user's context.

E. Personalized Information Presentation

In general, the output from the retrieved information is rarely the exact set of information desired by the user in response to a request [5]. Even if the retrieved information were precisely those that the user wanted, there exist few information retrieval systems that produce the personalized contents responding to the user's preference. On the other hand, the output from the retrieved information is limited by the output devices included in the user. For example, even when the preferable information is retrieved, if there is not an available output device to display, it may not be possible to present a satisfactory result to the user. Consequently, the information retrieval system fails apparently to present personalized information. Therefore, personalized information retrieval system requires that the preferred contents of the retrieved information should be reformatted based on a user's context such as transcoding technique [22,23]. In this regard, we can utilize any available display device like TV in ubiquitous computing-enabled home environment. In addition, it should support adequate output devices such as retinal display for personalized presentations [24,25]. However, we assumed that a user has or is able to acquire an appropriate output device. As explained in personalized information filter, we already have the filtered information based on the user's context. Thus, personalized information presentation is nothing but to display the filtered contents with an adequate manner.

F. Information Retrieval Measures

A viable information retrieval system must be effective in returning information in response to a user's request. While this generally means that most of the retrieved information in response to the user's request should be judged by the user to be appropriate to the information need, such a vague statement of effectiveness provides no solid basis for determining how good a given system is, or for comparing one retrieval system to another [4-6]. Thus, we employ *recall and precision* method to measure the effectiveness of personalized information retrieval system because it is the most widely used.

In recall and precision measurement, we can establish a 2 x 2 contingency table which shows how the information set is divided by both relevance and retrieval. Table I shows the relationships of these two measures.

TABLE I
CONTINGENCY TABLE FOR EVALUATING INFORMATION RETRIEVAL

	Retrieved	Not Retrieved	
Relevant	w	x	$n_1 = w + x$
Not Relevant	y	z	
	$n_2 = w + y$		$N = w + x + y + z$

In recall and precision method, precision is defined as the proportion of retrieved information that is relevant as follows.

$$P = \frac{w}{n_2} \tag{5}$$

Recall is defined as the proportion of relevant information that retrieved as follows.

$$R = \frac{w}{n_1} \tag{6}$$

Precision and recall are both bounded above by 1 (when $y = 0$ and $x = 0$, respectively), and below by 0 (when $w = 0$). With this evaluation method, we can approximately evaluate the performance of personalized information retrieval system.

IV. EXPERIMENTAL RESULTS

In this Section, we will demonstrate the effectiveness of the proposed framework with a practical example and a prototype application in ubiHome test-bed, where we use wear-UCAM as a fundamental development tool [2,18]. Figure 4 illustrates the experimental setup in ubiHome.

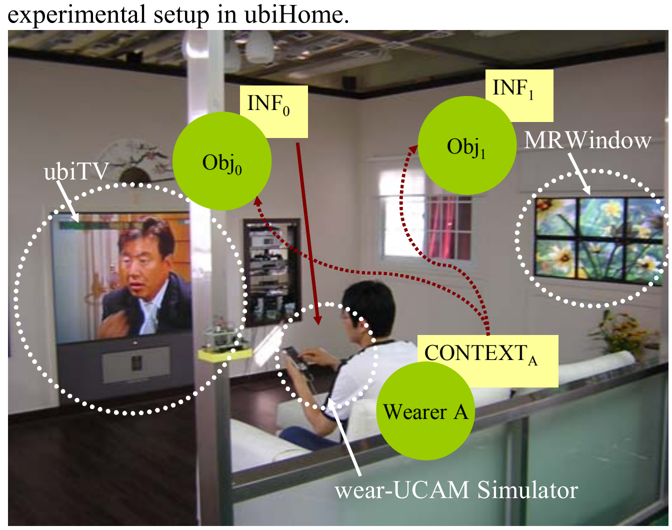


Fig. 4. An Experimental Setup for Personalized Information Retrieval System.

For the brevity of simplicity, we exploit only two objects, i.e. *ubiTV* and *MRWindow*, in ubiHome. Furthermore, the two objects have their own properties or information as the unified context format. Meantime, we acquire user’s context from the explicit user’s inputs, which is also represented as 5W1H format. As shown in Figure 4, the wear-UCAM simulator is used on behalf of a wearable computer, and it utilizes several sensors to acquire user’s preference. Although we did not take account of actual sensors to acquire user’s context for this experiment, it might be possible to exploit virtual sensors from wear-UCAM toolkit. Meanwhile, a general description of user’s context can be described as “*I want ubiTV service to be turned on if I am at home in the evening and my stress is relatively high*.” Regarding this description of the user’s context, we will show how to exploit user’s context in the proposed framework. Furthermore, we assumed that the user is interested in *when* and *what* terms only.

Given the description of user’s context, we can generate contextual query. Figure 5 illustrate some parts of user’s context in XML and its contextual query in the temporary data structure.

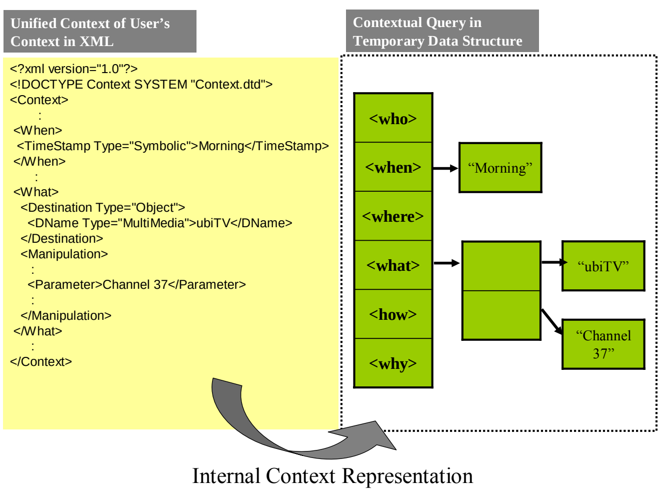


Fig. 5. An Example of User’s Context and Contextual Query.

As shown in Figure 5, a user’s context can be described in markup scheme model, XML. This markup scheme enables us to describe the user’s context with various attributes, and it is a well-defined document structure. Moreover, we can determine how many elements each term has in advance because of DTD (Document Type Definition) in the used XML for representing the unified context [14,26,27]. Therefore, we can keep the framework consistent from any contextual description model as well as keep user’s context exchangeable among context-aware applications. Similarly, we can also represent two objects as an object database.

Owing to the unified context representation of user’s context and object’s information as well as internal data structures for them, we can simply transform them to the weighted Boolean format. Regarding weights, we acquired weight of each term from user’s explicit input. First of all, each term of user’s context and object’s information is represented as a *0-1* vector. Second of all, we convert each term to a numerical value by the following strategy. If each element of terms in the *0-1* vector is corresponding to 1, we assign it to a vector form of ASCII values (string) because a string can be comparable. With this representation, we can transform user’s context to weighted Boolean query form. For the brevity of simplicity, we assumed that we are interested in *when* and *what* terms, where the weights of the two terms are 1 and rest of them are 0. Another assumption is that each weight of the terms is acquired from the user’s explicit input. Therefore, the transformed user’s context is

$$\mathbf{Q'} = <0, <\text{“Morning”}>, 0, <\text{“ubiTV”}, \text{“Channel 39”}>, 0, 0>,$$

where we present a string instead of a vector form of ASCII values. However, it must be noted that user’s context and object’s information are definitely coincident. Likewise, we can also achieve the transformed **Obj'** from the available objects. However, each term of objects reflects current status of them in order to deliver right information when the user pays attention to an object. It is often useless to deliver information when the user has lost his/her attention to the object [7]. The transformed objects are

$$\text{Obj}'_0 = <<\ldots>, <\text{“Morning”}>, <\ldots>, <\text{“ubiTV”}, \text{“Channel 37”}>, <\ldots>, <\ldots>>,$$
$$\text{Obj}'_1 = <<\ldots>, <\text{“Morning”}>, <\ldots>, <\text{“MRWindow”}, \text{“Flower”}>, <\ldots>, <\ldots>>.$$

where we have only presented when and what terms because the rest of terms are canceled out in the computation of the matching function.

After context transformation, we are able to select an object of interest and retrieve preferred information through personalized object filter and personalized information filter, respectively. In the selection of an object, we use the matching function as shown in (3). However, if q_i and obj_i include a set of strings, then the comparison of strings proceeds in advance. After string comparison, we just take a summation of each element. For example, if $q_3 = <\text{“ubiTV”}, \text{“Channel 37”}>$ and $obj_{0,3} = <\text{“ubiTV”}, \text{“Channel 37”}>$, then $q_3 \cdot obj_{0,3} = <1, 1> = 2$. On the other hand, $q_3 = <\text{“ubiTV”}, \text{“Channel 37”}>$ and $obj_{1,3} = <\text{“MRWindow”}, \text{“Flower”}>$, then $q_3 \cdot obj_{1,3} = <0, 0> = 0$. Another tactic in the computation of denominator in the matching function is to count the number of elements in each term as a numerical value. For example, $\Sigma(q_i)^2 = |0^2+1^2+0^2+2^2+0^2+0^2| = 5$. With this auxiliary computation method, each value of the matching function is evaluated as follows.

$$M(\text{Q}',\text{Obj}'_0) = 0.4472,\ M(\text{Q}',\text{Obj}'_1) = 0.1491$$

As a result, we can select *ubiTV* as the object of interest because it is larger than *MRWindow* object. However, the results of the matching function are not exactly what we expected because we treat the rest of terms in Obj'_0 and Obj'_1 as 1 in this computation. Thus, we need more investigation on context transform, especially a numeric conversion from a string. In the mean time, in the case of retrieving only one object from others, we do not have to consider *Th* value because we take an object with the largest value in $M(\mathbf{Q'},\mathbf{Obj'}_k)$. However, the threshold value *Th* is an important factor to determine the performance of information retrieval system.

With the retrieved object of interest, now we have to present appropriate contents only. In this application, the user wants to retrieve information of "*Channel 37*" in the morning among other properties. In this case, we simply retrieve "*Channel 37*" properties recursively from the object's information in object database. Therefore, users are able to retrieve personalized information only based on their own context.

Regarding time complexity of the proposed framework, we employ Big **W**(n) notation, especially the worst case. As we indicated, context transformation is a key feature in the proposed framework. However, it only requires at most $\mathbf{W}(m_{max})$, where m_{max} is simply determined by maximum number of element in terms. Meanwhile, the worst case of personalized object filter occurs when all the elements of each term in either **Q'** or **Obj'** have a set of strings. If there are m_1, m_2, m_3, m_4, m_5 and m_6 elements in each term, then worse case of time complexity is still $\mathbf{W}(m_{max})$ in the proposed framework. However, we can reduce time complexity by assigning the weights of irrelevant terms as 0. Thus, we can guarantee that personalized information retrieval algorithm is linearly dependent on the number of elements in context.

To evaluate the effectiveness of an information retrieval system, it is common to compute the precision-recall values. Thus, we conduct experiments with virtual data sets in order to evaluate overall performance of the proposed framework. The experimental conditions are as follows. We randomly generate 10 sample sets which represent objects' information as numerical values. Likewise, we generate a sample query set for user's context. The number of relevant objects is 5. Figure 6 illustrates the plotted graph of precision and recall.

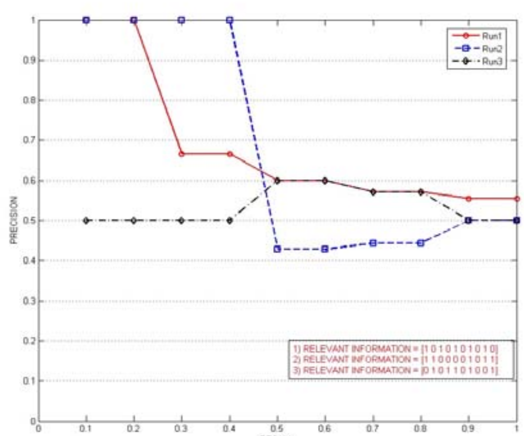


Fig. 6. Precision and Recall Graph of the Proposed Personalized Information Retrieval Framework.

Using figures based on precision and recall requires knowledge a priori of relevance, i.e., the answer set. However, this is not possible in general. Moreover, relevance is difficult to define satisfactory. In this regard, it is also hard to exploit precision and recall as a measurement of information retrieval system without enough experiences on this metrics. Thus, we briefly explain how to use it in practice. Typically, precision-recall values are computed at fixed recall levels from 0.1 to 1.0 in steps of 0.1 [1, 10].

Figure 7 shows similarity between query and collection of object, where we might determine a threshold value for information retrieval dynamically.

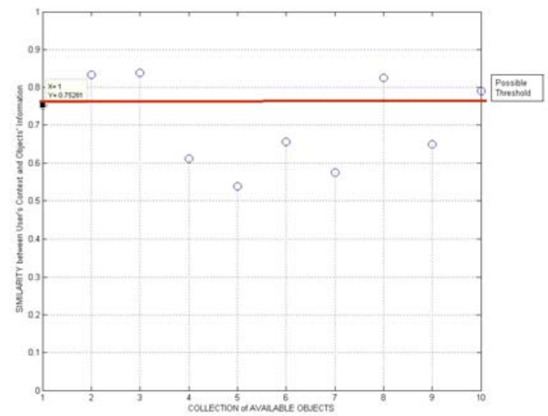


Fig. 7. Similarity Graph between User's Context and Object's Information.

Lastly, we would like to mention context transformation which transforms arbitrary contextual information to numerical values. While we were developing this framework, we found that if we transform context to any numerical expression, then we could apply many mathematical approaches to information retrieval systems. This explains why the previous research on information retrieval is mainly on a specific domain, where we can easily employ the pre-knowledge or predefined tables for numerical references on contextual information like Ontology. In this regard, we believe that context presentation scheme is fundamentally required, and its numerical transformation is also a corner stone for not only personalized information retrieval but also other context related research areas.

V. Discussion and Future Works

In this paper, we have proposed a framework of personalized information retrieval system for wearers in ubiquitous computing environment with an example of deterministic approach, a weighted Boolean. The proposed framework exploits user's context within the well defined components when the user is trying to retrieve information from objects in an environment. In particular, we transform user's context to the contextual query statement, which is represented as a weighted Boolean. Thus, the proposed framework enables us to retrieve only preferred information to any output device in the given environment. However, we still need more investigations on the following issues in order to realize personalized information retrieval system in the future computing environment. For instance, 1) how to transform contextual information of users and objects to numerical values simply but yet effectively, 2) how to deliver the retrieved information to the user appropriately i.e., where and when to show, and 3) how to apply time factor in user's preference because user's interests are varying on time, which apparently includes feedbacks from either the system or the user.

REFERENCES

[1] M. Weiser, "The Computer for the 21st Century," Scientific American, pp. 94-104, Sep. 1991.

[2] Seiie Jang, Choonsung Shin, Yoosoo Oh, and Woontack Woo, "Introduction of "UbiHome" Testbed," The first Korea/Japan Joint Workshop on Ubiquitous Computing & Networking Systems 2005 (ubiCNS2005), Proceeding CD, 2005.

[3] B.J.Rhodes, P.Maes, "Just-in-time information retrieval agents," IBM SYTEMS Journal, vol. 39, No. 3&4, pp.685-704, 2000.

[4] William B. Frakes, Ricardo Baeza-Yates, "Information Retrieval: data structures and algorithms," Prentice Hall, 1992.

[5] Korfhage, Robert R., "Information storage and retrieval," John Wiley&Sons, 1997.

[6] C. J. van Rijsbergen, "Information Retrieval," Butterworths, London, second edition, 1979.

[7] S.Julier, M.Lanzagorta, Y.Baillot, L.Rosenblum, S.Feiner, T.Höllerer, and S.Sestito, "Information filtering for mobile augmented reality". In Proc. ISAR '00 (Int.Symposium on Aug-mented Reality), pages 3-11, Munich,Germany, October 5-6 2000.

[8] Seth Insley, "Obstacle to General Purpose Augmented Reality," http://islab.oregonstate.edu/koc/ece399/f03/final/insley2.pdf.

[9] G.J.F. Jones, P.J. Brown., "Context-aware retrieval for ubiquitous computing environments," Invited paper in Mobile and ubiquitous information access, Springer Lecture Notes in Computer Science, Vol. 2954, pp. 227-243, 2004.

[10] Patrick A. S. Sinclair, Kirk Martinez, David E. Millard, and Mark J. Weal, "Augmented Reality as an Interface to Adaptive Hypermedia Systems", In New Review of Hypermedia and Multimedia, Special Issue on Hypermedia beyond the Desktop. Vol. 9, pp.117-136, 2003.

[11] Simon Julier, Yohan Baillot, Marco Lanzagorta, Dennis Brown, Lawrence Rosenblum, "BARS: Battlefield Augmented Reality System", NATO Symposium on Information Proc-essing Techniques for Military Systems, 9-11 October 2000, Istanbul, Turkey.

[12] B. Bell, S. Feiner, and T. Hollerer, "Information at a Glance," IEEE Computer Graphics and Applications, Vol. 22, No. 4, pp. 6-9, July/August 2002.

[13] A.K.Dey and G.D.Abowd, "Towards a better understanding of context and contextawareness," In the Workshop on the What, Who, Where, When and How of Context-Awareness, affiliated with the 2000 ACM Conference on Human Factors in Computer Systems. 2000.

[14] Albrecht Schmidt, "Implicit Human Computer Interaction Through Context," Personal Technologies Volume 4(2&3), June 2000, pp. 191-199, 2000.

[15] Seiie Jang, Woontack Woo, "Unified Context Describing User-Centric Situation: Who, Where, What, When, How and Why," The first Korea/Japan Joint Workshop on Ubiquitous Computing & Networking Systems 2005 (ubiCNS2005), Proceeding CD, 2005.

[16] Karen Henricksen,Jadwiga Indulska, and Andry Rakotonirainy, "Modeling Context Information in Pervasive Computing Systems," Proceedings of the First International Conference on Pervasive Computing, Lecture Notes In Computer Science, Vol. 2414, pp. 167-180, 2002.

[17] T. Strang, C. Linnhoff-Popien, "A Context Modeling Survey," First International Workshop on Advanced Context Modelling, Reasoning And Management UbiComp 2004, Nottingham, England, September 7, 2004.

[18] Dongpyo Hong and Woontack Woo, "wear-UCAM: A Toolkit for Wearable Computing," The first Korea/Japan Joint Workshop on Ubiquitous Computing & Networking Systems 2005 (ubiCNS2005), Proceeding CD, 2005.

[19] J. Choi, M. Kim, and V. V. Raghavan, "Adaptive feedback methods in an extended boolean model," In Proceedings of ACM SIGIR Workshop on Mathematical/Formal Methods in Information Retrieval, New Orleans, LA, Sept. 2001.

[20] Frei H.P. and Qiu Y., "Effectiveness of Weighted Searching in an Operational IR Environment," In Schriften zur Informationswissenschaft, Band 12, pages 41-51, UVK, Universitätsverlag, Konstanz, 1993.

[21] H. Chen, B. R. Schatz, C. Lin , "Concept Classification and Search on Internet Using Machine Learning and Parallel Computing Techniques," World Wide Web 4 Conference, Boston, MA, Pages 58-59, Poster Proceedings, December 11-14, 1995.

[22] R. Han, P. Bhagwat, "Dynamic Adaptation In an Image Transcoding Proxy For Mobile Web Browsing", IEEE Personal Communications Magazine, Dec. 1998, pp. 8-17.

[23] Claudio Pinhanez, "The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces," Proc. of Ubiquitous Computing 2001 (Ubicomp'01), Lecture Notes In Computer Science; Vol.2201, pp.315-331, Atlanta, Georgia, September 2001.

[24] http://www.microvision.com

[25] M. von Waldkirch, P. Lukowicz, G. Tröster, "Spectacle-Based Design of Wearable See-Through Display for Accommodation-Free Viewing," Pervasive Computing, pp.106-123, Vienna, Austria, 18.-23. April 2004.

[26] http://www.w3c.org/XML/

[27] http://java.sun.com/xml/jaxp/dist/1.1/docs/tutorial/overview/3_apis.html

Dongpyo Hong received his B.S. degree in Computer Engineering from Dong-A University, Busan, Korea, in 2001 and M.S. degree in Information and Communications from Gwangju Institute of Science and Technolgoy (GIST), Gwangju, Korea in 2004. Now he is a Ph.D student in GIST since 2004. Research Interest: HCI, Information Retrieval, Context-awareness, Vision based User Interface, Wearable computing, Ubiquitous computing, VR, and Entertainment Computing.

Yun-Kyung Park Yun-Kyung Park received her B.S. degrees in mathematic education from Korea University, Korea, in 1987. Since 1987, she has been working at Electronics and Telecommunications Research Institute, and she is a principal member of engineering staff of the Post-PC Platform Research Team. Her research interests include context-awareness, and sensor-networking.

Jeongwon Lee received his B.S. in Electrical Engineering, Ohio University, Ohio, USA, in 1989 and M.S. in Electrical Engineering, Univ. of Wisconsin, Madison, USA in 1991. He received his Ph.D. in Computer Science, Univ. of Southern California, California, USA. In 2002, as a research associate, he worked in Integrated Media Systems Center, California, USA. In 2002, as an Assistant Professor, he joined Digital Contents, Sejong University, Seoul, Korea. Research Interest: Computer Graphics, Augmented Reality, Virtual Reality, Computer Vision, Digital Contents, Computer Games, and Ubiquitous Computing.

Vladimir Shin received the B.Sc. degree and M.Sc. degree in applied mathematics from Moscow State Aviation Institute, Russia, in 1977 and 1979, respectively. In 1985 he received the Ph.D. degree in mathematics at the Institute of Control Science, Russian Academy of Sciences, Moscow. From 1984 to 1999 he was Head of the Statistical Methods Lab. at the Institute of Informatics Problems, Russian Academy of Sciences, Moscow. During 1995 he was Visiting Professor in the Mathematics Department at Korea Advanced Institute of Science and Technology. From 1999 to 2002 he was Member of Technical Staff at the Samsung Institute of Technology. He is currently an Associate Professor at Gwangju Institute of Science and Technology, South Korea. His research interests include estimation, filtering, tracking, data fusion, stochastic control, identification, and other multidimensional data processing. He has authored or coauthored over 50 papers in these fields.

Woontack Woo received his B.S. degree in EE from Kyungpook National University, Daegu, Korea, in 1989 and M.S. degree in EE from POSTECH, Pohang, Korea, in 1991. He received his Ph. D. in EE-Systems from University of Southern California, Los Angeles, USA. During 1999-2001, as an invited researcher, he worked for ATR, Kyoto, Japan. In 2001, as an Assistant Professor, he joined Gwangju Institute of Science and Technology (GIST), Gwangju, Korea and now at GIST he is leading U-VR Lab. Research Interest: 3D computer vision and its applications including attentive AR and mediated reality, HCI, affective sensing and context-aware for ubiquitous computing, etc.