

# INFORMATION SELECTION AND DELIVERY ALGORITHM FOR DELIVERING ADVERTISEMENTS SUITABLE FOR THE PEDESTRIANS PRESENT AT A PARTICULAR SITE

Koichi MATSUOKA<sup>1</sup>, Nobuo KATOU<sup>1</sup>, Shigeo DEJIMA<sup>2</sup>  
and Kazumasa TAKAMI<sup>1&2</sup>

<sup>1</sup> Graduate School of Engineering, Soka University, Tokyo, Japan  
e09d5201@soka.ac.jp

<sup>2</sup> Faculty of Engineering, Soka University, Tokyo, Japan  
k\_takami@soka.ac.jp

## **ABSTRACT**

*We consider a service that shows an advertisement suitable for the preferences of the pedestrians present at the site where the display is installed. In order to evaluate the effectiveness of this service by a simulation that reflects the reality well, we propose evaluation functions that take account of the line of sight of moving pedestrians. The evaluation functions incorporate two independent variables: Category and Producer. Based on this evaluation functions, we also propose an algorithm of multicasting an advertisement to different sites in order to deliver advertisements efficiently over a network while ensuring a high advertisement relevance rate. The evaluation results show that the proposed algorithm always scores a higher advertisement relevance rate than conventional advertising media. It is also shown that the proposed algorithm can reduce the delivery time by about 37%, and the volume of data delivered from the server by about 39% compared with unicast delivery.*

## **KEYWORDS**

*Ubiquitous Network, RFID, Delivering Advertisement, Information Selection and delivery algorithm*

## **1. INTRODUCTION**

True to its meaning, the word “ubiquitous” is now found everywhere. With a view to realizing a mature ubiquitous network society, great efforts are being made around the world to address a variety of issues involved, such as the establishment of social infrastructures, standards and security policy, to say nothing of the R&D on RFID technology, which is the key element in such a society. As a ubiquitous network matures, every object around us will be connected to the network either permanently or on an ad hoc basis, and many things and services impossible to implement today will become a reality. One such thing is identifying people’s preferences for different purposes. This is now attracting great interest. Such services are already provided for PCs. They identify the type of content the user is likely to take interest in based on, for example, his/her search record [1]. When everything around us is connected to the network, many services that make life easier will become possible by considering that what a person has is indicative of his/her preferences. There are many R&D activities on extracting people’s preference information based on information about the objects they use. This information is collected using RFID, which is now attracting a great deal of attention as a means of connecting things to the network [2]-[7].

There is also growing interest in how to deliver advertisements to people when there is a mature “ubiquitous network environment”. Since the preferences of individuals are becoming more and more diverse and personalized, “only now”, “only here”, and “only for you” are said to be key words in the 21<sup>st</sup> century marketing. There are already services that use the information about the location of an individual mobile phone user to provide a limited-time coupon of a shop in the vicinity of his/her location [8]. In such services, the direction of the flow information is necessarily from the surroundings of a person to a terminal he/she carries. Today, digital displays are increasingly installed as advertising media in public spaces in bigger cities. In public spaces, people often move in a group, such as a family, a group of friends, a loving couple, or a group of co-workers. Even if a person is alone, he/she often temporarily form a group with people who happen to have the same purpose as his/hers. They may be heading for the same office or for the same event site. Because of this tendency, it will be useful to extract the preferences of individuals present at a site, use this preference information to select an advertisement that is likely to interest most of those at the site, and deliver it to the digital display installed there. This advertising approach is more effective than conventional advertising media, such as posters and digital signage [9], that cannot change its advertisement to adjust to the people at the site where the advertisement is shown. If people are walking in a family or a group of friends, and if an advertisement suitable for them is shown, it may invigorate their communication. Or, those minority pedestrians whose preferences are not reflected in the advertisement shown because its selection is based on the preferences of the majority may be encouraged to take new interest in the advertised product. Although such a service is not provided today, it will become possible considering the great efforts being made for a ubiquitous network society and the progress in RFID technology. We have evaluated the advertisement viewing rate (hereafter referred to as the advertisement relevance rate) and data delivery efficiency for such a service using simulation.

This paper considers a service of delivering, to digital displays, an advertisement suitable for the pedestrians present at the site where each digital display is installed. To evaluate the effectiveness of this service using a simulation that reflects the reality well, we propose evaluation functions that take account of the line of sight of pedestrians. These evaluation functions incorporate two independent variables: Category value and Producer value. Both of these are elements of pedestrian’s preference information. To deliver advertisements efficiently over a network while ensuring a high advertisement relevance rate, we also propose an algorithm of multicasting an advertisement that is selected based on the proposed evaluation functions to each display. This paper is structured as follows. Section 2 introduces past studies on RFID-based services conceived for a ubiquitous network society. Section 3 outlines the service considered in this paper, the assumed environment and other assumptions. Section 4 identifies issues that need to be considered in simulating the assumed service. Section 5 proposes solutions to these issues: namely, the evaluation functions and the advertisement selection and delivery algorithm. Section 6 evaluates the proposed algorithm using simulation. Finally, Section 7 presents the conclusions and future issues.

## **2. RFID-BASED SERVICES IN A UBIQUITOUS NETWORK SOCIETY**

Great efforts are being made around the world to promote the use of RFID, which is an essential element of a mature ubiquitous network environment. Japan is the only country where studies and field trials are being made on the use of RFID tags attached to products already sold on the market, and where efforts are being made to promote the use of RFID and to build the necessary infrastructure. For example, the RFID tag embedded in a book may not be invalidated when the book is sold, but can be brought home and used for the management of book collection at home. The home appliance industry is studying the use of RFID tags for after-sales maintenance and recycling of products. In encouraging the use of RFID, it is important not only to use RFID for the management of products themselves, their distribution, and their history, but also to enable consumers to perceive the benefits of RFID in an environment in which RFID tags are normally

attached to products. Studies on how to extract necessary information from RFID tags, and activities to provide information extraction services are described below.

Reference [2] extracted people's preferences from their daily activities, and tried to enhance the functionality of an algorithm for providing a service suitable for individual users. Generally speaking, the objects a person puts in his/her room or carries are relevant not only to his/her individuality but also to his current work or things he/she is currently interested in. This paper proposed an algorithm for collecting people's behavior logs, such as objects they have put in their rooms or carried, and for extracting their preferences from these logs. It uses a smart *furoshiki* (wrapping cloth) [10], a cloth-type RFID reader, to collect people's behavior logs. This study assumed an environment in which RFID tags are normally attached to products sold on the market [11].

Reference [3] proposed a system that prevents people from losing their belongings. In this system, passive RFID tags are attached to objects people manage, and an RFID reader is installed in the environment. As a person walks past the RFID reader installed at the entrance to a room, the RFID tags attached the objects he/she carries are automatically read by the RFID reader and the identities of the objects are automatically found. The person does not need to be conscious of the system, and still can be alerted to anything he/she has left behind when he/she leaves the room. The system alerts the user by sending the information about the RFID tag attached to the object he/she has left behind using a short-distance wireless communication system based on mote [12].

Reference [5] proposes a system that determines the information a person needs from the objects he/she carries, and recommends this information to him/her. This system determines a person's situation by combining information about the objects he/she carries and his/her profile, which includes information about his/her location, the current time and the current temperature. It uses RFID passive tags and RFID readers to collect information about the objects he/she carries.

Reference [13] refers to a project that involved a field trial of using electronic tags to provide a futuristic shopping service. The field trial involved smart carts (shopping carts). They are used to provide promotional information without the consumers being aware of it, to provide smart shelves, which keep track of the inventory on individual shelves and provide product information to consumers, to provide fashion coordination information during trial fitting, and to allow express check-out, which expedites cashiers' work to reduce customers' waiting time. Another fashion coordination service is EPC Magic Mirror [14]. When a customer stands in front of a mirror holding a product, this service displays information about the product on the mirror.

### **3. SERVICE CONSIDERED IN THIS PAPER**

Digital displays are expected to be used for a variety of purposes in the coming ubiquitous network society. This paper considers a service of delivering, to a digital display, an advertisement suitable for the pedestrians present at the site where the display is installed. This service is targeted at areas in which there are many pedestrians and the roads are relatively wide. In such an area, advertising media are typically installed or pasted on pillars or walls on roads at intervals of 5 to 10 m, as is the case with a road that connects two relatively big train stations located close to each other, or an underground street that connects a train station to some landmark. Pedestrians with a variety of preferences walk in different directions. Although advertising media can be paper or electronic devices, this study assumes that only digital displays are used, that UHF-band RFID tags are normally attached to products sold on the market, and that RFID tags attached to objects pedestrians carry are read by RFID readers installed in the environment to obtain information about their preferences. EPCglobal has proposed to assign a GTIN (Global Trade Item Number) to each RFID tag [15]. JAN 13-digits

that comply with GTIN are assigned to products in Japan. It is possible to identify the category and producer (distributor, manufacturer, etc.) of an object from RFID information attached to it. In principle, there can be multiple Producers in one Category. In this study, the combination of Category and Producer values identified from an RFID tag carried by a person is considered to be his/her preference information.

Let us consider the relationship between the objects a person carries and his/her preferences, the communication distance of UHF-band RFID passive tags, and the attachment of these tags to objects. As is discussed in Reference [2], social psychology recognizes a strong relationship between the objects people carry and their preferences [16]. Since the objects people carry indicate the type of objects they are interested in and love, their preferences can be inferred from what they carry. Since UHF-band RFID passive tags do not have batteries, their communication distance is normally 3 to 5 m, but a communication distance of 10 m was achieved in Reference [17]. Reference [18] says that a system was developed in which UHF-band RFID passive tags were read from a distance of 180 m, and a maximum of 700 RFID tags were read per second. Since RFID tags may be used in a variety of environment and attached to a variety of objects, tags of difference shapes and characteristics are now being developed [19]. To realize a ubiquitous network society, it is necessary to study RFID technology with emphasis placed on creating new value that will make the life of consumers better, and to promote its use. We believe that the service considered in our study is suitable for the desired ubiquitous network society. The concept of the service is shown in Fig. 1.

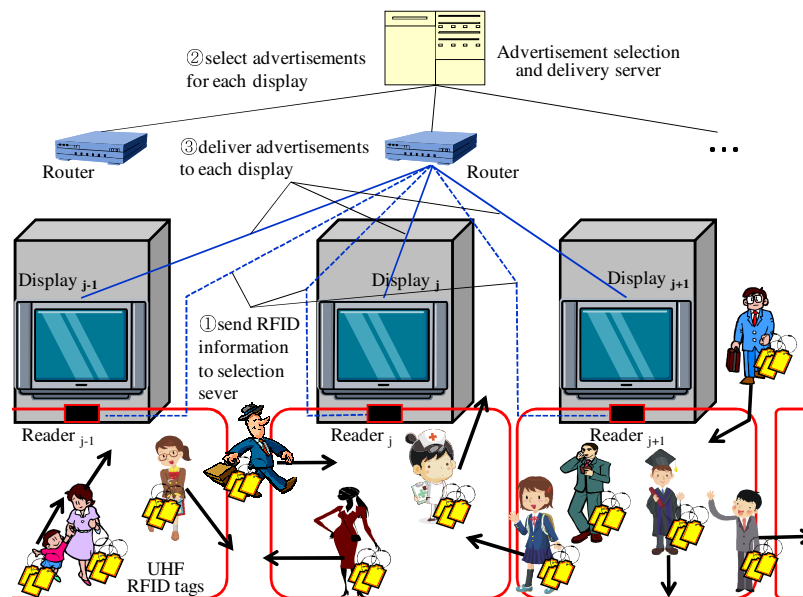


Figure 1. Assumed service

In the assumed environment,  $j$  displays are installed on pillars or walls, and an RFID reader is installed in the vicinity of each display. Hereafter, the display and the RFID reader at site  $j$  is expressed as  $D_j$  and  $R_j$ . Both the displays and RFID readers are connected to the advertisement selection and delivery server via a network. RFID information read by  $R_j$  is sent to the server periodically at intervals at which the RFID readers read tags. The server identifies the Category and Producer values from the RFID information it has received from all RFID readers. Based on the identified information, it selects an advertisement appropriate for the pedestrians at each site, and sends it to each display. It can be assumed that the selected advertisement will interest at least those pedestrians who carry objects with RFID tags that contain the same Category and Producer values as the advertisement, or those pedestrians who carry objects with RFID tags

that contain the same Category value as the advertisement. In the latter case, the Producer value of the object he carries is different from the Producer value of the advertisement, and so some pedestrians may be interested in the advertisement and some may not. It is not realistic to assume that all pedestrians are interested or no pedestrians are interested in the advertisement. This paper assumes that, if at least Category value of both the advertisement and the object a pedestrian carries is identical, he/she tends to be interested in the advertisement. To represent this tendency, we use the concept of viewing rate. Usually, the viewing rate used for TV programs. It is the estimated percentage of those TV owners at a certain site who have watched a certain program. The population of the viewing rate is not the households with their TVs turned on at a particular time but all the households that are the subjects of the viewing rate survey. In the same manner, in this study we define the viewing rate (or the “advertisement relevance rate”) as the estimated percentage of those pedestrians who carry an object with the same Category value as that of the advertisement among the pedestrians present at the site covered by  $R_j$  during a unit time, rather than the percentage of the pedestrians that carry an object with the same Category value as that of the advertisement shown on  $D_j$  among the pedestrians who are watching  $D_j$  during a unit time. For example, suppose that, at the time when an advertisement with Category value=A and Producer value=B is displayed, two-thirds of the pedestrians are found to have RFID tags with Category value=A (with any Producer value). In this case, we assume these two-thirds are interest in the advertisement, and thus, the advertisement relevance rate is  $2/3=66.7\%$ .

This study specifically aims to achieve two objectives. The first objective is that the advertisement relevance rate of our algorithm is higher than the case with fixed advertisements, such as posters and conventional digital signage. The second objective is that the selected advertisements are delivered efficiently so that the traffic load on the network will be reduced.

#### **4. SIMULATION-RELATED ISSUES OF THE ASSUMED SERVICE**

To perform a realistic simulation of the assumed service, it is necessary to take account of the line of sight of pedestrians as described in (1) below. To achieve the two objectives mentioned above, it is necessary to satisfy the condition described in (2) below.

##### (1) Line of sight of pedestrians

As shown in Fig. 2, it is assumed that any pedestrians within the site covered by RFID reader  $R_j$  can watch display  $D_j$ . Depending on the directions in which pedestrians move and on the interval at which displays are installed, pedestrians within the site covered by  $R_j$  can also watch  $D_{j-1}$ ,  $D_{j+1}$ ,  $D_{j+2}$ , etc. In addition, depending on the relative positions of  $D_{j-1}$  and  $D_j$ , a pedestrian within the site covered by  $R_{j-1}$  may find it easier to watch  $D_j$ . This means that the RFID information read by  $R_j$  should be used for the selection of an advertisement not only for  $D_j$  but also for other adjacent displays.

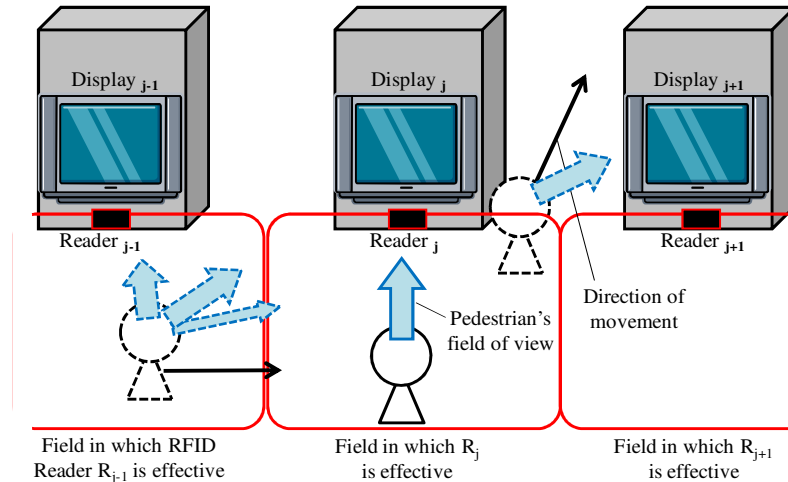


Figure 2. Effects of line of sight of pedestrians

## (2) Method of delivering an advertisement efficiently

In IP multicast communication, as many copies of one packet as necessary are generated within the network and delivered to as many destinations. If the same items of data are to be sent to multiple displays, IP multicast communication minimizes the amount of data that flow in the network, and achieves efficient data delivery. In the assumed service, two items of information, Category value and Producer value, are extracted from RFID information, and only one advertisement is selected. In general, the greater the range of value the two variables can take, the smaller the probability that the values of the two variables in one case take the exactly the same values of the two variables in another case. If the advertisement to be shown on each display is to be selected based on a combination of two variables contained in RFID information of objects carried by pedestrians, it is not very likely that the advertisements of more than two displays coincide with each other. This means that there are few cases in which IP multicast communication can be applied.

As mentioned in (1), the selection and delivery of advertisements can be made only by a server that can analyze the preferences of pedestrians at not only a single site but all the sites. In existing IP multicast communication, a multicast router groups together terminals to which the same items of data are to be delivered, makes copies of the data it has received from the data delivery server, and sends these copies to all the terminals belonging to a particular group. According to the current specification of IP multicast, whether a terminal joins, leaves, or stays in, a group is determined through a negotiation between the multicast router concerned and the terminals in the group. The same thing can be said of cache management. In other words, according to the current specification, even if a server is to send the same data to multiple terminals, the server needs to use unicast communication. As a result, more traffic load than is necessary is applied to the server and to the network, reducing the efficiency of data delivery. To solve this problem, it is necessary to develop a logic that enables the same advertisements to be shown on many displays so that data can be delivered efficiently using IP multicast. To provide the assumed service efficiently, the server needs to create GLs (group lists) for IP multicast and to notify the routers of them. It also needs to instruct routers to save data in their caches.

## 5. SOLUTIONS

This section proposes evaluation functions that take account of the line of sight of pedestrians mentioned in Section 4 (1), and an information selection and delivery algorithm that satisfy the condition mentioned in Section 4 (2).

## 5.1. Evaluation functions

To take account of the line of sight of pedestrians, we propose evaluation functions.  $\phi_1$  to  $\phi_6$ . Function, RFID (C, P, q)<sub>j</sub>, is defined as the total number of RFID tags that are read at the q-th reading cycle at site j and that have Category and Producer values of (C, P). Hereafter, the total number of RFID tags whose Category value is C but whose Producer value is arbitrary is expressed as RFID (C, \*, q)<sub>j</sub>. The total number of RFID tags whose Producer value is P but whose Category value is arbitrary is expressed as RFID (\*, P, q)<sub>j</sub>. It is to be noted that there is a one-to-many relationship between Category values and Producer values. For example, when Category value=A, there can be Producer value=a to z. However, in this study, we treat the two parameters as being independent of each other. Accordingly, Category and Producer values in  $\phi_1$  to  $\phi_6$  are treated as being independent of each other. The aim is to enable the proposed information selection and delivery algorithm to satisfy the two objectives of our study. As was mentioned in Section 4 (2), if an advertisement is to be selected based on both Category and Producer values, the advertisements selected for different displays become greatly varied, reducing the efficiency of advertisement delivery. To avoid such a situation, an evaluation function for Category and another for Producer need to be independent of each other in the information selection and delivery algorithm.

### 5.1.1. Evaluation functions that take account of the line of sight of pedestrians

One of the ways to reflect RFID information read by  $R_j$  to the selection of advertisement shown on  $D_k$  is to determine the direction in which each pedestrian is moving by referring to his/her RFID log. Once the direction of a pedestrian's movement is determined, it is possible to reflect his/her preferences to the advertisements shown on displays that stand in the direction of his/her movement. However, it is practically impossible to determine the direction of each pedestrian's movement with 100% accuracy. As is the case with the viewing rates of TV programs, even if the direction of a pedestrian's movement is determined correctly, which display he/she is watching cannot be determined with a high level of accuracy. To solve these problems, we propose a method of reflecting RFID information read at one site to the selection of advertisements to be shown on displays at other sites. Specifically, we propose evaluation functions,  $\phi_1$  and  $\phi_4$ , shown below in order to reflect the RFID information read by  $R_0$  to the selection of the advertisement shown on  $D_k$ . Here, it is assumed that there are RFID readers  $R_0$  to  $R_{TAIL}$  in the assumed environment.

$$\begin{aligned}\phi_1 &= \text{RFID}(C, *, q)_j \times (1/2)^{|j-k|} \\ \phi_4 &= \text{RFID}(*, P, q)_j \times (1/2)^{|j-k|} \quad (0 < j \pm k < \text{TAIL})\end{aligned}$$

Suppose that  $D_j$  stands in the north of a pedestrian at the site where  $D_j$  can be watched. Broadly speaking, the line of his/her sight is either east ( $D_{j+1}$ ) or west ( $D_{j-1}$ ). If the expectation of him/her watching  $D_j$  is 1, the expectation of him/her watching  $D_{j-1}$  and that of watching  $D_{j+1}$  are both 1/2. We also consider  $D_{j-k}$  and  $D_{j+k}$ . We assume that k displays are located adjacent to  $D_j$ , and that the expectation of the pedestrian watching  $D_{j-k}$  and that of watching  $D_{j+k}$  are both  $(1/2)^{|j-k|}$ .

### 5.1.2. Evaluation functions that take account of changes in the user's preferences

Evaluation functions,  $\phi_1$  and  $\phi_4$ , work in such a way that the selection of the advertisement to be shown on  $D_j$  is influenced by the information read by readers in other sites, in particular, readers in adjacent sites, such as the sites where  $R_{j-1}$  and  $R_{j+1}$  are located. There would be no problem if the preferences of the pedestrians in all sites are stable and uniform. However, this is an unrealistic assumption. In an environment in which pedestrians with different preferences move around freely, it is highly possible that the preferences identified at the site of  $R_j$  is overly influenced by the preferences identified in adjacent sites. In such a case,  $\phi_1$  and  $\phi_4$  may lead to a selection of an advertisement that is not the best choice for  $D_j$ . To solve this problem, we reduce

the influence of  $\phi_1$  and  $\phi_4$  by taking account of the preferences at the site of  $R_j$  in the previous reading cycle. For this purpose, we propose the following evaluation functions:

$$\phi_2 = \text{RFID}(C, *, q)_j + \text{RFID}(C, *, q - 1)_j \times (1/2)$$

$$\phi_5 = \text{RFID}(*, P, q)_j + \text{RFID}(*, P, q - 1)_j \times (1/2)$$

Evaluation function  $\phi_1$  dictates that the selection of the advertisement to be shown on an adjacent display is influenced by  $\text{RFID}(C, *, q)_j \times (1/2)$ . To reduce this influence, one half of  $\text{RFID}(C, *, q - 1)_j$  is added to evaluation function  $\phi_3$ . The same is also applied to evaluation function  $\phi_4$ .

### 5.1.3. Evaluation functions that cope with cases where the scores of different Category values are the same

Evaluation functions,  $\phi_1$ ,  $\phi_2$ ,  $\phi_4$  and  $\phi_5$ , allow an advertisement to be selected taking account of the viewing field of pedestrians and changes in their preferences. However, the application of these evaluation functions can lead to a situation in which multiple sets of preferences get the same score. To avoid this situation, we consider the ratio of the number of cases where a certain Category value occurs to the total number of RFID tags read in the assumed environment, and a similar ratio for Producer. To take account of these ratios, we propose the following evaluation functions:

$$\phi_3 = \text{RFID}(C, *, q)_j / \text{All Tags}$$

$$\phi_6 = \text{RFID}(*, P, q)_j / \text{All Tags}$$

where All Tags is the number of RFID tags read in the assumed environment.

Evaluation function,  $\phi_3$ , is effective in selecting an advertisement using  $\phi_1$  and  $\phi_2$  in cases where multiple Category values get the same score. Similarly, evaluation function,  $\phi_6$ , is effective with regard to Producer values.

### 5.1.4. Effectiveness of each evaluation function

We consider the situation shown in Fig. 3 to explain the effectiveness of each evaluation function. For simplicity, we consider only three sites that are adjacent to each other:  $R_{j-1}$ ,  $R_j$  and  $R_{j+1}$ . We consider only two Category values: “music” and “book.” The numbers at each site in the figure indicate the number of tags of a particular Category value at that site. Let us look at the case where  $q=1$  (the first reading cycle) in the figure. At the site of  $R_j$ , the number of tags for Category value=“music” is the same as that for Category value=“book”. However, since the preferences of the pedestrians at the adjacent sites who are assumed to be watching  $D_j$  are reflected, the score “music” gets is  $(60+50+40=)$  150 while the score “book” gets is  $(60+30+20=)$  110. In this way, even in cases where the simple comparison of the numbers of tags of different Category values does not lead to a definitive selection of an advertisement for a site, the use of  $\phi_1$ , which takes the line of sight of pedestrians into consideration, can lead to the selection of a suitable advertisement.

Next, we look at Case 1 ( $q=2$ ) in the figure. At the site of  $R_j$ , the number of tags with Category value=“music” is 40 while that with Category value=“book” is 60. The application of  $\phi_1$  gives the score of 140 to “music” and 130 to “book.” This result reflects the fact that the application of  $\phi_1$  to the site of  $R_j$  results in its scores being excessively influenced by the strong preferences to music at the site of  $R_{j-1}$ . If this result is used and as a result, an advertisement suitable for music is selected for  $D_j$ , the advertisement relevance rate would be lower than the case an advertisement suitable for book is selected for  $D_j$ . However, if  $\phi_2$  is applied to take account of the situation in the previous reading cycle, the score for Category value=“music” becomes  $(40+30=)$  70 while that for Category value=“book” becomes  $(60+30=)$  90. If this is added to the score calculated using  $\phi_1$ , the score for Category value=“music” becomes  $(140+70=)$  210 while that for Category value=“book” becomes  $(130+90=)$  220. If an advertisement is selected based



on these scores, the advertisement relevance rate of Category value="book" will go up. In this way,  $\varphi_2$  is effective in cases where taking the line of sight of pedestrians into consideration lowers the advertisement relevance rate.

Next, let us look at Case2 ( $q= 2$ ) in the figure. As in Case 1,  $\varphi_2$  is incorporated to prevent the preferences of adjacent sites from excessively influencing the selection of an advertisement for  $D_j$  as a result of the application of  $\varphi_1$ . In this case, the score of Category value="music" and that of Category value="book" are both 220. In this way, the applications of both  $\varphi_1$  and  $\varphi_2$  can lead to a situation in which  $t$  Category values get the same score. However, if  $\varphi_3$  is applied, the score for Category value="music" and that for Category value="book" becomes  $(40/460 \doteq) 0.09$  and  $(60/460 \doteq) 0.13$ , respectively. Here, 460 is the total number of tags in all the sites. This time, the score for Category value="music" exceeds that for Category value="book". If this result is used to select an advertisement, the lowering of the advertisement relevance rate can be avoided as was the case in Case 1. In this way,  $\varphi_3$  is effective in cases where multiple Category values get the same score.

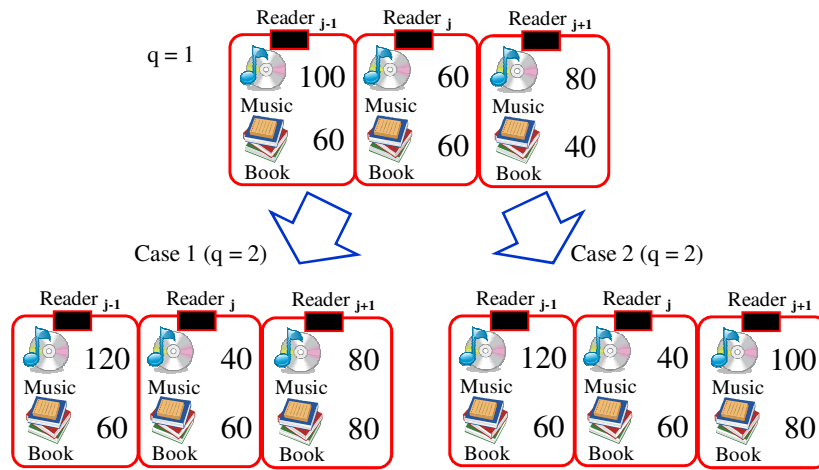


Figure 3. Different cases of pedestrians with different preferences

## 5.2. Advertisement selection and delivery algorithm

This section proposes the advertisement selection and delivery algorithm referred to in Section 4 (2) based on the evaluation functions proposed above. This algorithm assumes that there is a database in place that refers to JAN to identify Category and Producer values of RFID information. It was mentioned earlier that IP multicast communication must be used if we are to achieve efficient delivery of advertisements, that the server that analyzes RFID information collected by all the readers in the assumed environment must create GLs (group lists), and instruct routers to saving data in their caches if the line of sight of pedestrians is to be taken into consideration. This paper proposes multicast routers that receive GLs and instructions of saving data in their caches from the server, and use IP multicast communication where possible. Such multicast routers are assumed to be in place. The processing sequence of the proposed advertisement selection and delivery algorithm is described below:

- Step 1: Initialize each evaluation function
- Step 2: Collect RFID information from each  $R_j$  until a unit time has elapsed
- Step 3: After a unit time has elapsed, compare the collected RFID information with data in the database to identify Category and Producer values
- Step 4: Calculate  $\varphi_1+\varphi_2+\varphi_3$  for the site of each  $R_j$ .

- Step 5: Calculate  $\phi_4 + \phi_5 + \phi_6$  for the site of each  $R_j$ .
- Step 6: Select the Category value whose score is the highest in Step 4 for the site of each  $R_j$ .
- Step 7: Create a GL of displays for which the same Category value has been selected for the site of each  $R_j$ .
- Step 8: Select the Producer value whose score is the highest in Step 5 from among the RFID information whose Category value is the one selected in Step 6 for the site of each  $R_j$ . If the advertisement that has earned the highest score is not registered, select the one with a Producer value that has earned the second highest score in the same Category value. If the advertisement selected in Step 6 is not registered, search for the one with the Category value that has earned the second highest score.
- Step 9: Check the cache  
 If the needed advertisement is in the cache, deliver it from the cache, and go back to Step 2.  
 If the needed advertisement is not in the cache, determine whether the advertisement should be saved in the cache, deliver the advertisement selected in Step 8 along with GL to each router, and go to Step 2.

To achieve efficient IP multicast communication, the algorithm shown in Steps 7 to 9 is applied instead of creating a GL after an advertisement has been selected based on the Category and Producer values. In other words, when the server selects the Category value with the highest score for the site of each  $R_j$ , it creates a GL of the displays to which the advertisement is to be delivered simultaneously at that time. It can be expected that the number of displays that receive the same advertisement is large enough to achieve efficient multicast communication. To enable the advertisement that is delivered the most frequently to be saved in the cache in each multicast router, the advertisement with the highest score of the Producer value is selected in high priority from among the advertisements of the Category value selected after the creation of the GL, and is saved in the cache. This is expected to increase the number of opportunities to use the cache. As was mentioned in Section 5.1, unless Category and Producer values are treated independently of each other in the calculation of evaluation functions, the number of opportunities to use IP multicast communication and to use the caches would be smaller. Figure 4 shows the sequence described above of selecting an advertisement and creating a GL.

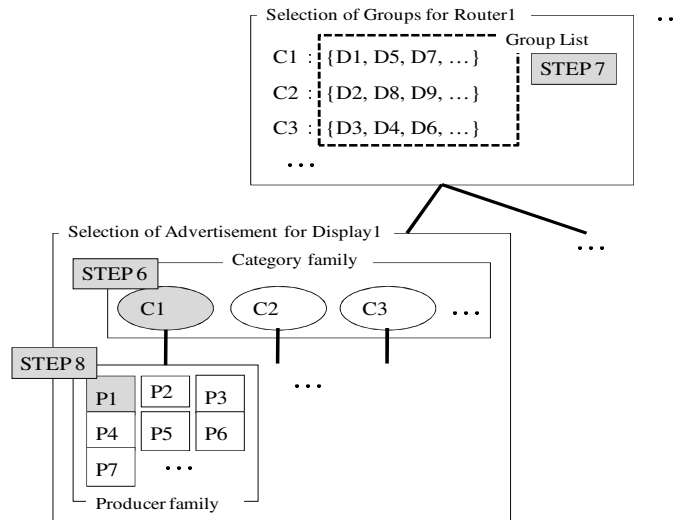


Figure 4. Sequence of selecting an advertisement and creating GLs

## **6. EVALUATIONS**

The proposed advertisement selection and delivery algorithm aims to satisfy two objectives: (1) to achieve a higher advertisement relevance rate than conventional advertising media, such as fixed posters and digital signage, and (2) to deliver data efficiently. To demonstrate the effectiveness of the proposed algorithm, this section compares this algorithm with conventional advertising media in terms of the advertisement relevance rate, and also compares the proposed multicast communication with unicast communication in terms of the data delivery time and the volume of data delivered.

### **6.1. Evaluation of the advertisement relevance rate**

Although the advertisements shown in digital signage are changed according the time of day, they are not adjusted to the pedestrians currently present at the site of the display. Therefore, digital signage can be considered to be a version of a fixed poster with the displayed information changed at long intervals. For this reason, digital signage is considered to be one of conventional fixed advertising media in this paper.

#### **6.1.1. Simulation model**

##### (1) Assumptions for the model

We assume that a person carries 10 to 15 objects with RFID tags attached to them. These objects may include about 5 items of clothing (a jacket and/or a shirt, and a pair of trousers), shoes, and bags, and 5 to 10 items of less essential objects, such as a music player and cosmetic items. The assumption that a person carries 5 to 10 less essential items is based on the “Survey on Objects Carried When People Go Out (excluding mobile phones and cash)”, which was published on February 8, 2005 by Inforplant (currently “Yahoo Value Insight”) [20]. This survey reports that men carry 4.8 items and women 8.4 items on average when they go out. The number of Category values and that of Producer values of advertisements are both considered to be 100. Although the number of Category values used in shopping sites or auction sites are about 300, some of them overlap in advertisements, and some are not suitable for advertisements. When these are excluded, 100 Category values can cover all types of products or advertisements. In most shopping sites, the number of Producer values in each Category value is about 30 at most. Therefore, 100 Producer values can cover all cases.

##### (2) Simulation model and simulation method

Figure 5 shows the simulation model. It is assumed that 30 RFID readers and 30 displays are installed in the assumed environment. Pedestrians may move from any direction to any direction. A person may abruptly turn back or change directions at any point. The pedestrians may enter or leave the site at random. At one instance, some pedestrians may enter a site from different directions and some may leave the site in different directions. The pedestrians may have different preferences. Therefore, to simulate the collection of RFID information, we generate a JAN-compliant 13-digit number per unit time as an input to each RFID reader. One RFID reader is associated with one display on a one-to-one basis. The generated numbers are input to the proposed algorithm, which selects the advertisement to be displayed on each display.

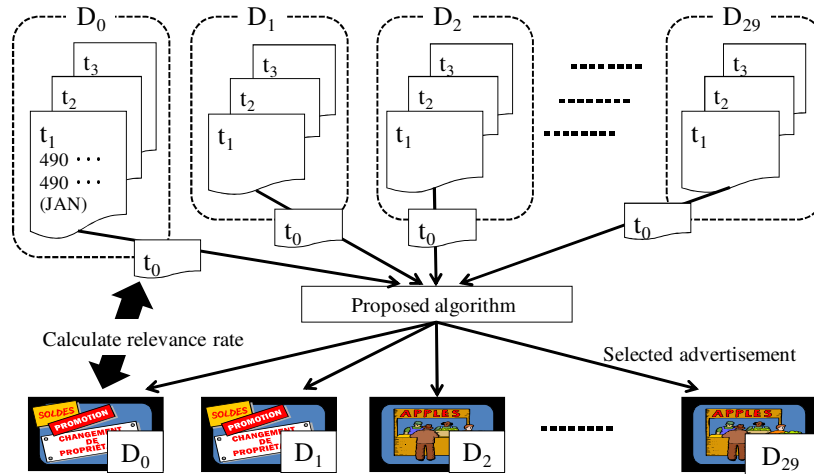


Figure 5. Simulation model

The simulation was conducted as follows. A text file used to generate a JAN-compliant random number is generated for each RFID reader per unit time. It is assumed that the number of people remains unchanged during one round of simulation. One file signifies the RFID information read by one RFID reader during unit time,  $t$ . The generated files are input to the algorithm at each elapse of unit time, and an advertisement is selected for each display. The advertisement relevance rate is calculated from the selected advertisements and the RFID information input to the proposed algorithm for each unit time. The average advertisement relevance rate for the entire environment is considered to be the advertisement relevance rate for the number of people within the environment. The number of people within the environment is the number of people at the site of  $R_j$  multiplied by the number of readers within the environment ( $=30$ ).

### 6.1.2. Evaluation results

Figure 6 shows the advertisement relevance rate for different numbers of people within the environment. The horizontal axis shows the number of people, and the vertical axis shows the advertisement relevance rate. Advertisement A is the preferred advertisement for 20 Category values representing clothing, shoes, bags and accessories. Advertisement B is the preferred advertisement for 80 Category values that represent less essential objects. Fixed advertisements A and B in Fig. 6 are these advertisements. The reason that the advertisement relevance rate of these fixed advertisements is as high as slightly above 10% is that the simulation model assumes that each person carries 10 to 15 objects. The probability that the RFID information generated at random matches a certain Category value is 10/100 to 15/100, where 100 is the number of Category values.

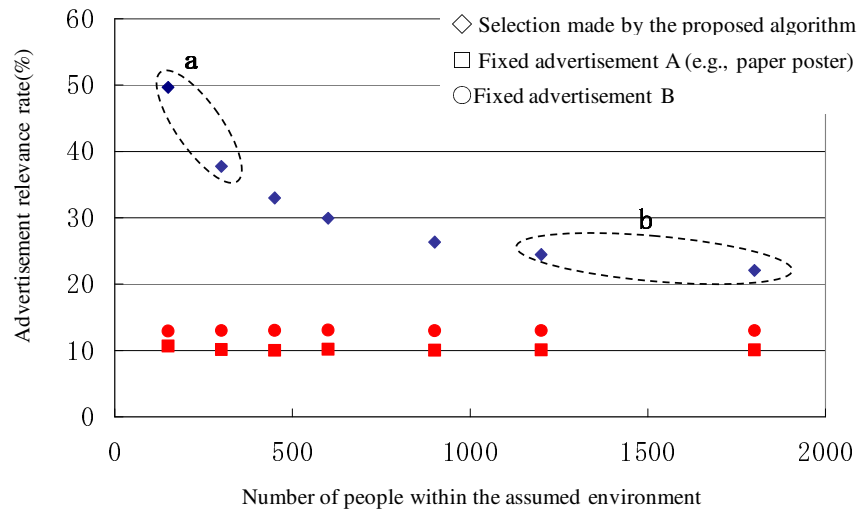


Figure 6. Advertisement relevance rate vs. number of people within the assumed environment

At ellipse a in Fig. 6, the number of people is 150. There are 5 people in the site covered by  $R_j$ . So, there can be 50 to 70 RFID tags in the site covered by  $R_j$ . This means some of the 100 Category values do not exist in this site. This means that the density of each Category value is relatively high. This explains why the advertisement relevance rate is relatively high. In contrast, at ellipse b, the number of people is 1800. There are 60 people in the site covered by  $R_j$ . So, there can be 600 to 900 RFID tags in the site covered by  $R_j$ . This means that there can be 6 to 9 RFID tags for each Category value. Consequently, the density of each Category value is more or less uniform. The advertisement relevance rate in this case is lower than the case where the density of each Category value is high. Still, the advertisement relevance rate achieved by the proposed algorithm is always higher than that achieved by fixed advertisements. It can be concluded that the proposed algorithm is particularly effective in cases where the number of people within the assumed environment is small, and that it is more effective than fixed advertisements even when the number of people is large. Figure 7 shows the variance of the advertisement relevance rate for the proposed algorithm. It shows that the variance becomes smaller as the number of people increases. This means that, as the number of people increases, the proposed algorithm can select advertisements at a more stable advertisement relevance rate.

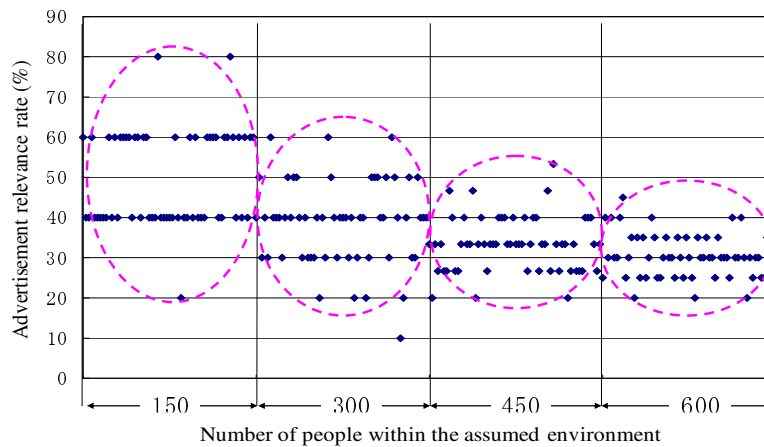


Figure 7. Variance of the advertisement relevance rate vs. number of people within the assumed environment

## 6.2. Evaluations of data delivery time and the volume of data

Since the proposed algorithm aims to use IP multicast communication as much as possible, the data delivery and the volume of data delivered was evaluated by comparing them with those when unicast communication is used.

### 6.2.1. Configuration of the evaluation system

To reflect a possible actual environment, the evaluation system is made up of one server and three multicast routers. Each router is connected with 30 displays and 30 RFID readers. The communication speed in the network is 100 Mbps, the most common speed in current optical communication. The size of a video clip is 2 MB, which is an approximate size of a 5-second-long CM. Each router has a cache with a capacity of 60 MB, which is about 12% of all the data delivered.

### 6.2.2. Simulation model

Figure 8 shows the delivery sequence assumed for evaluating the proposed algorithm in terms of delivery time and the volume of data delivered. In an actual network, communication delays vary depending on the loads on servers and routers and on how heavily transmission links are used. To take account of these, we simulated network delays by applying a random number in each step of delivery. The range of the random number is the theoretical delay time plus/minus 2.5 ms. Specifically, the delay time shown in Fig. 8 was used in each step. For ①, we chose 5 to 8 ms, which is slightly longer than the delay time for ping, because there are numerous, albeit small, items of RFID information. The delay time at ② and ④ can vary depending on the processing power of each server and router. We chose 1 to 2 ms for ② and ④ so that this delay time will not influence the evaluation results. For ③, we chose 3 to 6 ms, which is equivalent to the delay time for ping, because the size of a GL notification data is small. For ⑤, we chose 19 to 24 ms, which is the theoretical value for data transmission (video size / actual communication speed  $\times$  1000) plus/minus 2.5 ms. We added 18 to 22 ms to ①, ③ and ⑤ to account for a network delay.

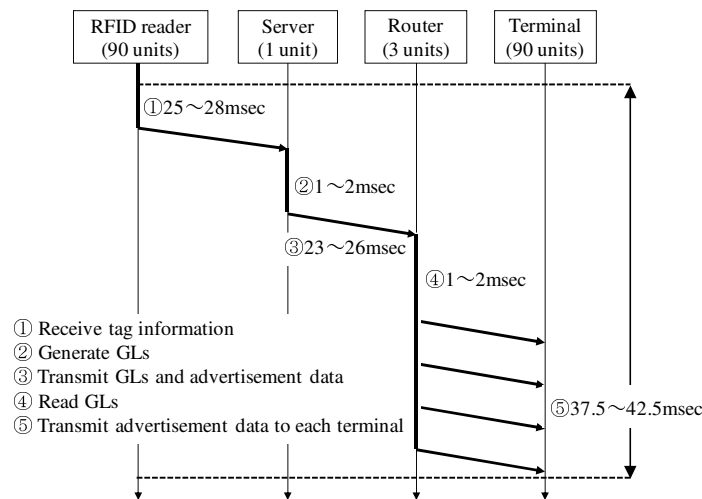


Figure 8. Advertisement delivery sequence

### 6.2.3. Evaluation results

Figures 9 and 10 show the simulation results for delivery time and the volume of data, respectively. The delivery time achieved by the proposed algorithm is shorter than that achieved by unicast communication by about 37%, and the volume of data delivered by the proposed algorithm is smaller than that by unicast communication by about 39%. Both the delivery time

and the volume of data vary because of the variation in the overlapping of Category values extracted from RFID information.

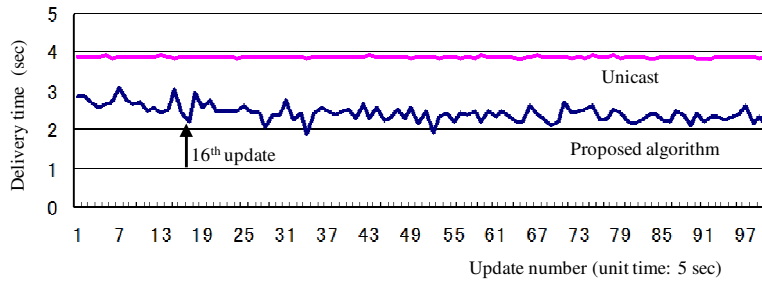


Figure 9. Deliver time vs. update number

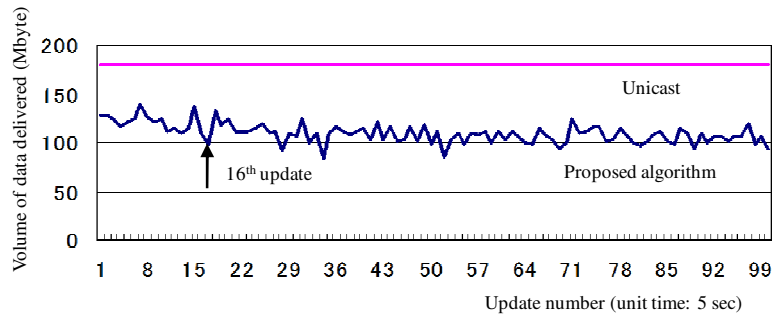


Figure 10. Volume of data delivered vs. update number

Figure 11 show the number of GLs for each GL update number (the sequence at which the GL was updated). The number of registrations in the figure indicates the number of GLs. The more the data to be delivered to displays overlap (Category values overlap), the smaller the volume of data delivered by the server according to the generated GL, and the shorter the delivery time becomes. For example, let us look at the 16<sup>th</sup> update, where the number of GLs is large. Figures 9 and 10 show that the delivery time is shorter and the volume of data delivered is smaller at the 16<sup>th</sup> update. These results indicate that the proposed algorithm actively generates GLs, and results in efficient data delivery.

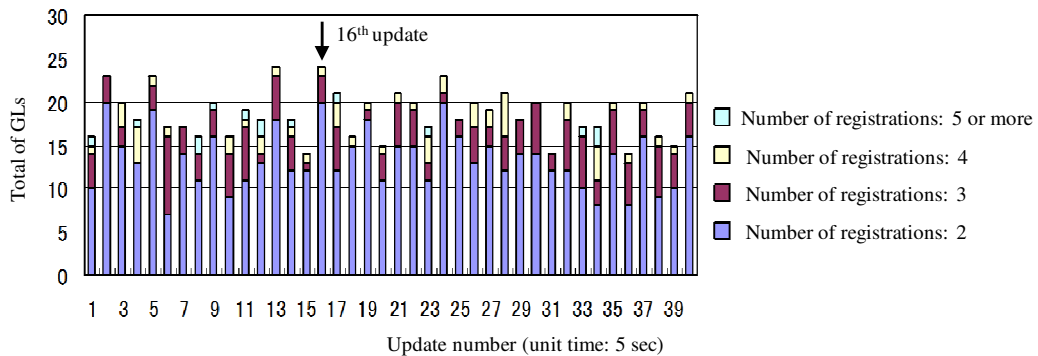


Figure 11. Number of GLs vs. update number

Figure 12 shows the percentage at which the caches are used. After the proposed advertisement selection and delivery algorithm has generated GLs during the processing of selecting Category values, it selects the Producer value with the highest score within the assumed environment, and saves the advertisement corresponding to that Producer value in the cache of each router in order to enhance the delivery efficiency. It can be expected that this advertisement will be selected relatively frequently. Figure 12 shows that, after the percentage of using caches reaches 12%, which is the target percentage of using caches, the percentage often drops below 12%. This indicates that the advertisements saved in caches are actually selected, resulting in efficient delivery. The average percentage of using caches is about 11%, which is slighter below the target percentage of 12%. This indicates that the proposed algorithm uses caches effectively even when the communication is unidirectional as is the case with the advertisement delivery service.

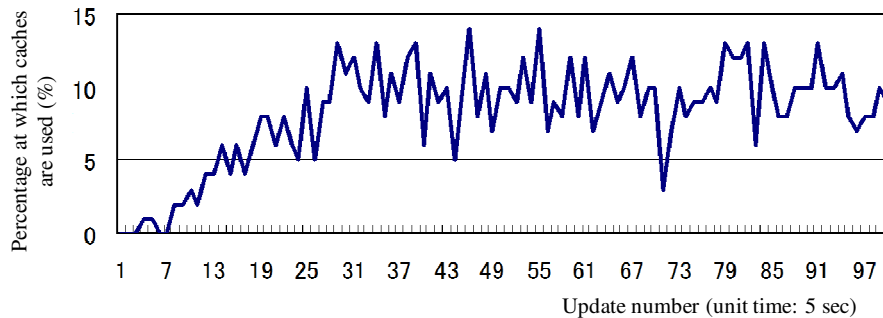


Figure 12. Percentage at which caches are used vs. update number

## 7. CONCLUSIONS

Activities to realize a ubiquitous network society include the use of digital displays installed in public spaces. This paper has considered a service in which an advertisement suitable for the pedestrians present near a display is delivered to that display. It is assumed that UHF-band RFID tags are normally attached to products sold on the market, and that the preferences of pedestrians can be identified from UHF-band RFID tags ready by RFID readers installed in the assumed environment. We have proposed evaluation functions  $\phi_1$  to  $\phi_6$  that take account of the line of sight of pedestrians, and have explained the effectiveness of these functions. We have then proposed an information selection and delivery algorithm that is based on the proposed evaluation functions. We have evaluated the proposed algorithm in terms of the advertisement relevance rate, data delivery time and the volume of data delivered. The evaluation results show that the proposed algorithm is particularly effective when the number of people within the environment is about 150, and that, even if the number of people within the environment is larger, the proposed algorithm still gives a higher advertisement relevance rate than conventional fixed advertising media. It is also shown that the proposed algorithm generates GLs actively, leading to a frequent use of multicast delivery. As a result, the delivery time is reduced by about 37% compared with unicast communication, and the volume of data is reduced by about 39%. In addition, it has been shown that the proposed algorithm can use caches effectively even in unidirectional communication, as is the case with the advertisement delivery service. Since these results are consistent, it can be said that the proposed information selection and delivery algorithm achieves both a high advertisement relevance rate and efficient data delivery.

Our future work includes the evaluation of the time consumed by a series of processes involved in the assumed service, from reading of RFID tags by RFID readers to the display of advertisements, and the identification of implemental issues for a commercial service.



## REFERENCES

- [1] <http://www.google.co.jp/>
- [2] Kei Suzuki, Ryo Ohsawa, Masayuki Iwai, Kazunori Takashio, Hideyuki Tokuda, "An Algorithm for Extracting User's Preference based on Historical Usage of User's Objects," IPSJ SIG Technical Report, UBI, 2006(116), pp.85-91, 2006.
- [3] Gaetano Borriello, Waylon Brunette, Matthew Hall, Carl Hartung, and Cameron Tangney, "Reminding about tagged objects using passive rfids," UbiComp, pp.36-53, 2004.
- [4] Kenneth Fishkin, Bing Jiang, Matthai Philipose, and Sumit Roy, "I sense a disturbance in the force: Long-range detection of interactions with rfid-tagged objects," UbiComp, pp.268-282, 2004.
- [5] Naoki Shimizu, Riko Yagiu, and Masashi Saito, "A study of information services based on personal belongings," AINA'05: Proceedings of the 19th International Conference on Advanced Information Networking and Applications, pp.523-528, 2005.
- [6] David C. Wyld, 24-Karat Protection: RFID and Retail Jewelry Marketing, International Journal of UbiComp (IJU), pp 1-14, Vol 1, Num 1, January 2010.
- [7] Vijayakumar.G.Dhas, Ramanathan Muthukaruppan, Konguvel Balakrishnan and Rajarajan Ganesan, Effective Load Balancing with Power Conservation in RFID, International Journal of UbiComp (IJU), pp 13-21, Vol.1, No.4, October 2010.
- [8] IMANARA! <<http://imanara.jp/index.html>> (2010.12.17).
- [9] DSC: Digital Signage Consortium <<http://www.digital-signage.jp/>> (2010.12.17).
- [10] Ryo Ohsawa, Masayuki Iwai, Takuya Imaeda, Kei Suzuki, Takuro Yonezawa, Kazunori Takashio, and Hideyuki Tokuda, "Smart-furoshiki: A sensorized fabrics supporting office activities," UbiComp, 2006.
- [11] Auto-id.japan. <<http://www.autoidlab.jp/>> (2010.12.17).
- [12] Crossbow. Wireless Sensor Network mote. < <http://www.xbow.jp/motemica.html> > (2010.12.17).
- [13] Ministry of Economy, Trade and Industry, "Future Store Project Japan" < <http://www.meti.go.jp/press/20051108001/20051108001.html> > (2010.12.17).
- [14] RFID JOURNAL < <http://www.rfidjournal.com/article/view/4429/>> (2010.12.17).
- [15] EPCglobal, "EPC Tag Data Standard (TDS)Standard" < <http://www.epcglobalinc.org/standards/tds/>> (2010.12.17).
- [16] Russell W Belk, "Possessions and the extended self," Journal of Consumer Research, Vol.15, pp68-139, 1988.
- [17] Y. KAWAKITA O. NAKAMURA J. MURAI, "Permissible Link Quality for RFID Anti-Collision in a Practical Environment," IEICE TRANS. on Comm. Vol.E91-B No.5 pp.1480-1489, 2008.
- [18] Mojix Takes Passive UHF RFID to a New Level, RFID JOURNAL < <http://www.rfidjournal.com/article/articleview/4019/1/1/> > (2010.12.17).
- [19] Shigeru Hashimoto, Koken Yamamoto, Yoshiyasu Sugimura, "RFID Tag Equipment Supporting Today's Ubiquitous Society," FUJITSU.57, 4, pp274-279 (05, 2006)
- [20] Investigation about The property at the time of going out (except for a mobile phone and cash), < <http://overturekeyword.blog.shinobi.jp/Entry/11057/> > (2010.12.17).

## Authors

Koichi Matsuoka

He received a master's degree in Faculty of Engineering from Soka University, Japan in 2009 and he is now in Soka University as a Ph.D. candidate of Soka University. He studies about a ubiquitous network services.



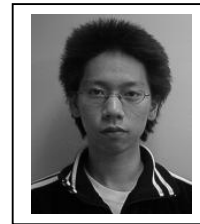
Nobuo Katoh

He received a master's degree in Faculty of Engineering from Soka University, Japan in 2007. He studied about ubiquitous services with RFID technology.



Shigeo Dejima

He received a bachelor's degree in Faculty of Engineering from Soka University, Japan in 2009. He studied about IP multicast method for advertising delivery.



Kazumasa TAKAMI

He received the B.S. and M.S. degrees in electrical engineering from Shizuoka University in 1977, 1979, respectively. In 1997, he received Ph.D. degree from Osaka University. He joined Musashino Electrical Communication Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Musashino, Japan in 1979. After that, he was involved in various projects, including the development of inter-networking systems between telephone and packet switching, the research into multi-media packet switching architecture, communication software for ATM switching system and the service-specification techniques for telecommunication software. Also, he was engaged in the development of Advanced IN systems, personalized IP services platform and next generation IP network architecture. Since April 2004, he has been a professor in Department of Information System Science, Faculty of Engineering, Soka University. He is a member of the IEICE, IPSJ, and a senior member of the IEEE.

