Recommendation-Aware Smartphone Sensing System

Mu-Yen Chen[1], Ming-Ni Wu[1], Chia-Chen Chen*[2], Young-Long Chen[3] and Hsien-En Lin[1]

[1] Department of Information Management
National Taichung University of Science and Technology
Taichung, Taiwan
[2] Department of Management Information Systems
National Chung Hsing University
Taichung, Taiwan
*emily@nchu.edu.tw
[3] Department of Computer Science and Information Engineering
National Taichung University of Science and Technology
Taichung, Taiwan

ABSTRACT

The context-aware concept is to reduce the gap between users and information systems so that the information systems actively get to understand users' context and demand and in return provide users with better experience. This study integrates the concept of context-aware with association algorithms to establish the context-aware recommendation systems (CARS). The CARS contains three modules and provides the product recommendations for users with their smartphone. First, the simple RSSI Indoor localization module (SRILM) locates the user position and detects the context information surrounding around users. Second, the Apriori recommendation module (ARM) provides effective recommended product information for users through association rules mining. The appropriate product information can be received effectiveness and greatly enhanced the recommendation service.

Keywords: Context-aware, smartphone sensing, recommendation service.

1. Introduction

In the concept of enterprise management, enterprises need to identify their target markets and create demand to bring profits. Therefore, in the present era of information explosion, how to quickly and accurately meet the target demand, timely supply the demand, and cope with planned marketing strategy to bring enterprise profits is one key concern of enterprises nowadays. With the rapid growth of data mining approach, the association rules derived from the analyses of huge transaction records can be applied to individual transaction records, so individual consumer marketing strategies can be developed [1]. Adopting the aforementioned method of developing marketing strategies will increase the overall sales revenue. In addition, the method of developing marketing strategies is more informed and effective and will bring the enterprises greater profits.

According to Kowatsch & Maass (2010), they found that consumers, in physical shopping environments, are strongly dependent on information present in such physical environments [2]. In other words, consumers are strongly dependent on the relevant information provided by the retailer regarding the commodities. Therefore, this study developed a recommendation system in physical shopping environments based on the concept of context-aware with Radio Frequency Identification (RFID) and smartphone as the medium. Consumers will use the handheld smartphone to read the RFID tag on commodities via RFID reader; smartphone will then provide and recommend relevant information of the commodities to consumers as to enhance the visibility of information for consumers in a physical shopping environment. It is thus to be concluded that one smartphone recommendation system which functions as the medium of information provider via a context-aware environment indeed influences consumer behavior.

2. Materials and methods

2.1 RFID Positioning Mechanism

Recently, there are many positioning algorithms or wireless mechanisms for RFID extensively proposed [3]. Hightower and Borriello (2000) presented the famous SpotON Indoor Positioning Technology, this technology is for detection of unknown objects with RFID-based indoor positioning technology [4]. The experiment used the RFID reader as the core and established center coordinates and massively spread tags on coordinates to record the received signal strength indication (RSSI) on each tag at its corresponding coordinates. The record of these RSSI was then used as data to analyze and infer the positions of the tags. The main localization mechanism are including Signal strength (SS), Time of arrival (TOA), Time difference of arrival (TDOA) [5], and Angle of arrival (AOA) [6]. According to the aforementioned positioning methods, there are varieties of positioning methods that are based on RSSI and each method has distinctive advantages and disadvantages. Take LANDMARC positioning mechanism [7] as an example. It is one indicative indoor positioning technology among current studies, but it is, at the same time, one with high costs and requires massive arrangement. The cause of the shortcoming is that LANDMARC positioning requires all three pieces of hardware equipment including RFID Reader, Reference Tag, and Tracking Tag to maintain indoor positioning accuracy. However, in the real physical environment, the arrangement of RFID equipment requires maintenance and raises environmental impact issues [8,9]; such issues need to be overcome one by one. Therefore, the goal is to establish context-aware in a physical shopping environment, in other words, it is to establish a RFID-positioning-technology-based consumer recognition function so that the system can provide relevant information based on a consumer's position so as to achieve ubiquitous computing concept.

2.2 Data Mining Technique

Generally speaking, data mining can be interpreted as knowledge discovery in database (KDD). In other words, knowledge can be extracted from massive data stored in one large database; such large databases can be on-line databases of data warehouses. The so-called "knowledge" indeed refers to some rules. Grupe and Owrang (1995) also claimed data mining is to dig out new facts which are still unknown to experts from the existing database [10]. After that, Fayyad et al. (1996) also discussed data mining is one step of the knowledge discovery process; to obtain special patterns from a large amount of data through data analysis algorithms [11]. Cabena et al. (1997) also explained data mining is the process of extracting previously unknown information with the highest relevance from databases, in order to use it in the decision-making process [12]. Recently, Han and Kamber (2001) summarized that data mining can extract the knowledge from large amount of data and discover interesting patterns [13]. In the traditional data mining approach, there are four popular methodologies, including data classification, data association, data clustering, and sequential patterning mining [13]. In 1994, Agrawal and Srikan presented the Apriori algorithm to handle the association rule mining [1]. It also called the market basket analysis and the goal is to investigate the possible relationship between different products from complex and huge transaction datasets. In this study, the proposed system will apply the association rule technique into transaction database and obtain the purchasing behavior or personal shopping behavior. Then, the system can send the related product information to users' smartphone and enhance the recommendation service quality.

2.3 Sensing Technique in Purchasing

According to the study of Kowatsch & Maass [2], it is found that consumers, in physical shopping environments, are strongly dependent on information present in such physical environments. In other words, consumers are strongly dependent on the relevant information provided by the retailer regarding the commodities. Therefore, the study developed one Mobile Recommendation Agent (MRA), a recommendation system in physical shopping environments based on the concept of context-awareness with Radio Frequency Identification (RFID) and Personal Digital Assistant (PDA) as the medium. Consumers will use the handheld PDA to read the RFID tag on commodities via RFID reader; PDA will then provide and recommend relevant information of the commodities to consumers as to enhance the

visibility of information for consumers in a physical shopping environment. Such concept is identical to the definition of Context-Awareness.

3. Architectural framework of the proposed system

3.1 System Modules

The Context-aware Recommendation Systems (CARS) are mainly made of Simple RSSI Indoor Localization Module (SRILM), Apriori recommendation module (ARM) as shown in Figure 1. The system deployments mainly are multi-device platform as well as smart handheld devices, and the integration of RFID technology. In addition, multi-device platform will be connected to the external Application Program Interface (API) of community media. Hence consumers, that is to say the users, can quickly access or share shopping information. The following provides a brief description of the conceptual system architecture in Section 3.2; and the primary user context are defined as follow.
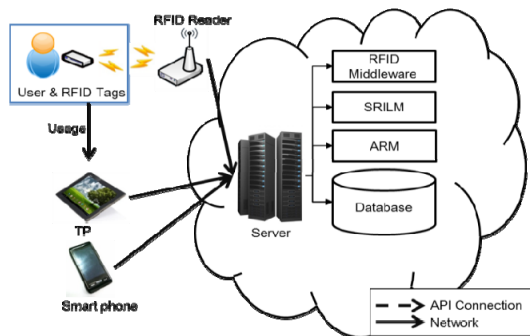

Figure 1. Context-aware recommendation system.

Provide consumers with ubiquitous recomm-ended context of shopping information via computing technology. Consumers will identify himself/herself through multi-platform interface operating system and RFID labels. Meanwhile, he/she will simulate that he/she is in the entity shopping mall. Deploy RFID readers is equipped in shopping centers with positioning consumer location mechanism. Software as a Service (SaaS) incorporates information such as locations of consumers, consumer expenditure record information. Through Apriori relational rule algorithm, compute the

product information of a desired consumer, and provide the information to consumer via multi-platform interface. In this way, consumers can obtain omnipresent commodity recommending information via entity shopping environment.

Provide consumers with ubiquitous recomm-ended context of shopping information via computing technology. Consumers will identify himself/herself through multi-platform interface operating system and RFID labels. Meanwhile, he/she will simulate that he/she is in the entity shopping mall. Deploy RFID readers is equipped in shopping centers with positioning consumer location mechanism. Software as a Service (SaaS) incorporates information such as locations of consumers, consumer expenditure record information. Through Apriori relational rule algorithm, compute the product information of a desired consumer, and provide the information to consumer via multi-platform interface. In this way, consumers can obtain omnipresent commodity recommending information via entity shopping environment.

3.2 System Architecture

Main structures of the proposed system are divided into five layers, namely, RFID layer, multi-platform layer, cloud layer, enterprise layer. Each layer provides its own services. Through internet connection, services of each layer are serial to form the architectural framework of the proposed system, as shown in Figure. 2.

3.2.1 Physical Layer-RFID Layer

In RFID layer, it contains RFID readers and RFID tags. RFID tag represents user identification and RFID readers are distributed in the physical store environment. Consumers, via putting the RFID tags close to the reader, let the RFID reader reads and hence the system recognizes his/her identity. RFID readers are distributed at different regions among stores and they can detect the positions of consumers. When RFID readers in the region recognize user identities, then the system will perform recommendation calculations pertaining to merchandises in the region. And let consumers located in that area obtain recommended merchandise information through handheld devices like smartphone.
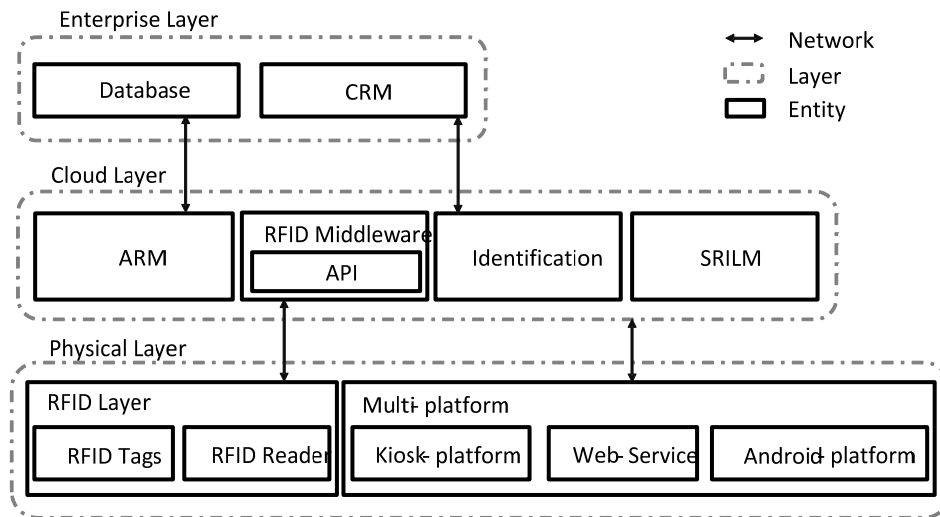
Figure 2. CARS deployment.

3.2.2 Physical Layer-Multi-platform

This layer will build a user interface that conforms to the multi-platform. It is connected to the cloud layer via the internet, including smartphone using Android 4.0 operating system. As to the smartphone, after user is registered as a member, the system will provide the member a membership serial number and a password. Users have to login to the system via the membership serial number and password. After users login their smartphone, if an RFID reader installed in the store region senses the RFID membership card with the users, through a middleware connecting to the database for comparison and confirmation, the system will present product information to the smartphone via wireless network.

3.2.3 Cloud Layer

Cloud layer consists Apriori algorithm recommendations, RFID localization, identification, and RFID middleware. To use the Apriori algorithm, it can extract recommendations rule from transaction records database and connect to enterprise level as the basis for recommendation computation. The RFID localization determines user's location based on RSSI signal strength and Simple RSSI Indoor Localization methods. The RFID membership identification is based on the comparisons of returning results of RFID middleware and

membership database to confirm membership. Finally, the RFID middleware provides responses to read-in result API for RFID readers deployed in various areas.

3.2.4 Enterprise Layer

Enterprise layer is for the business intelligent purpose of this system. It stores the membership database and transaction database and provides the value-added service to cloud layer, such as using Apriori algorithm to compute the association rule for product recommendations, or providing member database to identify the user membership. Hence, the enterprise layer is mainly used as a data source provider.

4. Experimental Design and Operational Scenario

4.1 Simple RSSI Indoor Localization Module (SRILM)

According to Hightower and Borriello (2000) [4] and our previous research [14], the SRILM module of the proposed model is primary using SpotON indoor positioning technology. The core operation of SRILM module is using RSSI as the kernel positioning technology, and also measuring signal strength between RFID reader and tag to calculate the possible distance to the RFID reader as shown in Figure 3. The goal is to calculate the correct

positioning of users. In the previous researches, the direction factor was not considered and calculated in original RSSI methodology. Nowadays, the active RFID reader and tag have the advantages on detecting distance range than passive ones. Therefore, this study define its reading range in straightforward method. When the user holds the RFID tag and wintin the detection range of active RFID reader, then this study will conclude the RFID tag in within the reading range. In addition, this method can satisfy and implement the simple positioning requirements of consumer location in the physcial store.
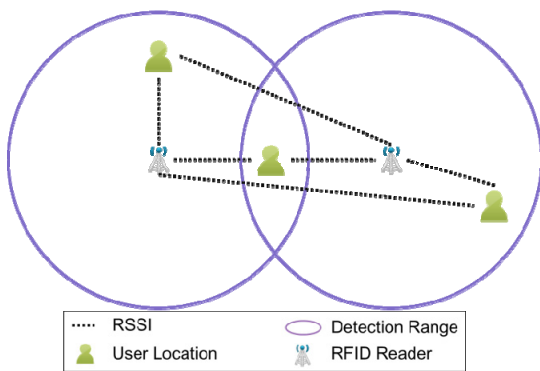

Figure 3. Illustration of SRILM.

The scenario illustration of SRILM algorithm is shown in Figure 4. In this study, the experimentla environment supposes that exists n users in the RFID-based shopping environment,

and User_Matrix $= [U_1, U_2, \cdots, U_n]$, each user U_n holds a RFID tag representing his/her personal identity represented as T_i. Besides, the RFID reader m is represented as Readers_Matrix $= [R_1, R_2, \cdots, R_m]$. Finally, $Readers_j$ has a certain signal coverage and is represented as *Area*. When RFID tag T_i. is detected by RFID reader $Readers_j$, then the signal is traced back, it refers that the RFID tag T_i. is within signal coverage of RFID reader $Readers_j$. Namely, location of T_i. is Area. In this study, the simple RFID indoor localzation method is described in the follows [14].

Step 1. When one RFID tag T_i has been detected by the RFID reader in the shopping store, then the signal intensity of this RFID tag can be represented as a vector $T_i = [Rs_1, Rs_2, \cdots, Rs_m]$, where Rs_1 is the signal strength value of RFID reader R_1 to connect to T_i. The largest value in the T_i, it means the specific RFID reader is the closest to the RFID tag.

Step 2. Find the maximum value $\max(T_i)$ of signal intensity in RFID tag T_i. In the other way, it represents the corresponding RFID reader covering range when T_i position is at $\max(T_i)$.

Step 3. If $\max(T_i)$ has the same maximum value at different locations, it implies that RFID tag T_i is located under specific coverage of multiple RFID readers, and then T_i is located at the junction of more than one area.
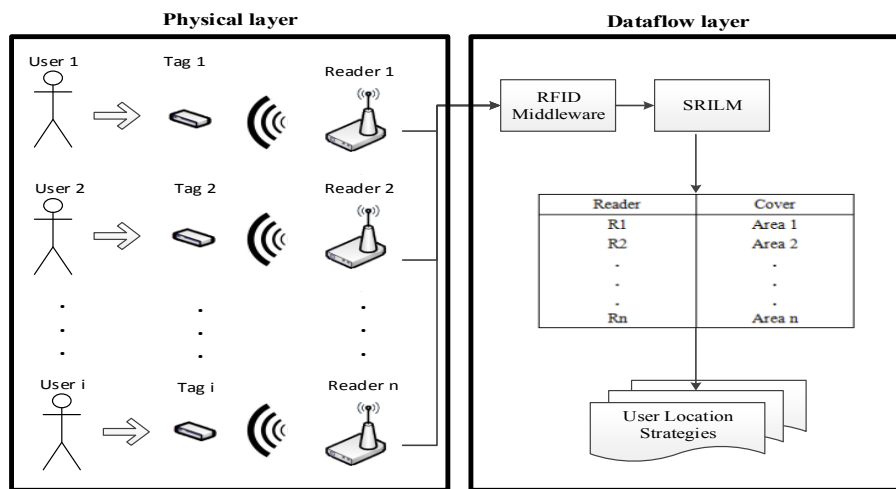

Figure 4. SRILM operation process.

4.2 Apriori Recommendation Module (ARM)

The system built in this study adopted association rule to extract the product recommendations. In association rule approach, the Apriori algorithm proposed by Agrawal et al. (1994) is a well-known data mining algorithm [1]. The recommendation process of this study is shown in Figure 5. When the consumer stays in the physical shopping environment, the system can identify the position of the consumer through held RFID tags, and then using signal strength as the basis for calculating the user coordinate with SRILM module. In addition, the use of ARM modules, it can identify the large itemsets and decide whether are the recommended itemsets or not from consumer transaction records. In addition, a collection of product recommended itemsets all have a product category position data fields in the database, so the proposed system can identify products in the physical environment position.

The algorithm is a step-by-step approach to identify the relationship of the database items. When the most items appear and the number of occurrences is the highest group, the combination of this group is the main rules of these data with important parameters as the maximum itemset size, minimum support and minimum probability.

Probability that is conditional probability refers to the accuracy of this rule. The equation of counting the possibility that B happened in condition A is as Equation 1 below.

$$\text{Confidence}(A \rightarrow B) = \frac{P(A \cup B)}{P(A)} \qquad (1)$$

Apriori algorithm computing steps:

Step 1. Scan all transaction records to join the combination of all itemsets into single objects.

Step 2. The frequency of a rule appear in all transactions does not meet the minimum standard of support, and the itemset fails to meet support standard will be pruned to retain itemsets that meet the standard and called as Candidate Large itemsets).

Step 3. Process Step 2 on itemsets that generate two item number and filter the itemsets fail to reach the minimum support standard.

Step 4. Repeat the above steps until the maximum rule items appear to be +1. Based on the generated itemset frequency division, association rule is found, and then have the rules fail to meet the minimum confidence filtered. The main processes are join and prune.
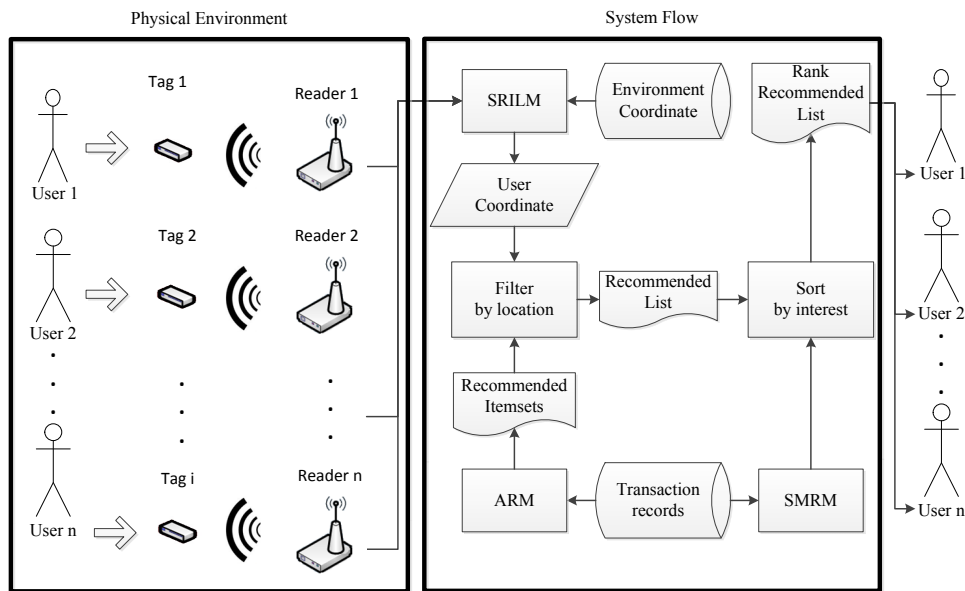

Figure 5. Recommendation of ARM.

If consumers stay longer in certain product category area, it can be assumed that consumers maybe want to buy these products, so system can position the surrounding consumer product categories as the basis. Then, from the collection of recommended items to filter out the items that contain the same set of category, then decide the recommended list of recommended list. Finally, the SMRM modules will reference the recommended list of products and calculate the magnitude of the products of concern, and finally sort the degree of concern descending order of products and become a ranked list of recommendations through the SNSs database API.

By the results of Apriori algorithm operations in this system, businesses can take advantage of eliminating inventory management system to collect the purchasing history of consumers and identify valuable recommendation product information from such information by Apriori algorithm of association rules in data mining. Such recommendation information can be downloaded by consumers from smartphone APP to provide them a more complete experience.

```
00   // System operational relationships between
01   three modules
02   Event
03   Define AT_r as All Transaction records
04   Define AR_s as itemsets recommend by ARM
05   Define LR_s as itemsets filter by SRILM
06   Define RR_l as list sort by SMRM
07   CarsRecommendationProcess(Event);
08   {
09   AR_s = ARM(AT_r);
12   LR_s = SRILM(AR_s);
11   RR_l = SMRM(LR_s);
12
13   Return RR_l
14   }
15
16   // Using ARM to find out the AR_s
17   C_k: Candidate itemset of size k
18   L_k: frequent itemset of size k
19   L_1 = {frequent items};
20   for (k = 1; L_k!= Ø; k++) do begin
21   C_k+1= candidates generated from L_k;
22   For each transaction t in database do
23   increment the count of all candidates in
24   C_k+1 that are contained in t
25   L_k+1= candidates in C_k+1 with
26   min_support
27   End
28   Return ∪_k L_k;
29
30   // Using SRILM to filter out theAR_s and
31   generate the LR_s
32   FilterByLocation(AR_s)
33   {
34   For each role in AR_s //role refers to the
35   recommendation rule of AR_s
36   {
37   For each item in role
38   {
39   If (item's shelf equal to user's Correct shelf)
40   // item refers to the products of recomme-
41   ndation rule, shelf refers to the physical
42   location of item
43   {Role add to LR_s
44   Break;}
45   }
46   }
47   }
```

Table 1. Pseudo code of the proposed CARS system.

4.3 Scenario Design

In order to test the result of positioning and to verify the accuracy of system positioning, scenario experiment is set to simulate consumer in physical environment. The experimental environment is a space with size of 10m*10m. The plane coordinate diagram of the environment is added to the system as the display screen of result of positioning. Readers are deployed in the four corners of the experimental environment and labeled Reader 1 to 4 respectively. Each coordinate is placed with a Tag wherein the Tag represents the position of consumer. During the positioning process, we define the Tag as the tracing label to calculate the coordinates of the Tag in this simulated physical environment, as shown in Figure 6.
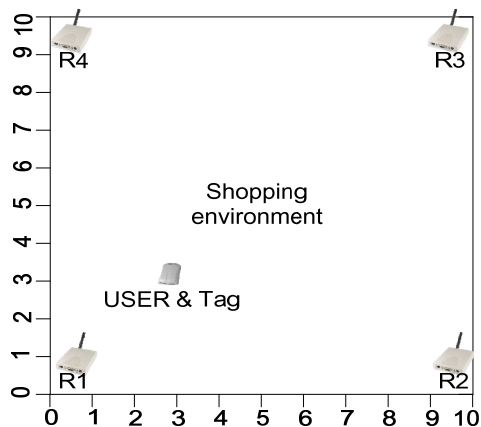

Figure 6. SRILM positioning mechanism operation screen.

This study simulates the scenario of consumer in physical environment. With the signal strength RSSI as the basis and via RFID and the RFID tag sensor held by consumer, SRILM module calculation is performed to find out the user coordinate.

In order to support the possibility of the proposed system, the hardware and software requirements of this experimental design are described in the follows. The operation system of Android Smartphone is Android 2.2 or above version. The operation system of database server is adopted Windows 7 64-bit version. In addition, the database software is Microsoft SQL Server2008 R2. Finally, the recommendation service is used the Apriori algorithm which is provided in Business

Intelligence Analysis Services in SQL Server 2008 R2. The operational scenario will be conducted for two parts. First, the location identification in different detection range of the RFID readers by SRILM positioning mechanism will be evaluated. Second, products will be recommended by using Apriori algorithm to users so that they can browse them on their smartphones.

4.4 CARS Performance Analysis

This aims of this study is to analyze the effectiveness of CARS recommended performance. First, this study simulated the consumer spending in the shopping environment, and collected their shopping nad transaction records in order to facilitate the association rules mining from consumers' shopping history. In this study, we collected 394 samples from consumer transaction records. In the traditional data mining approach, for example, association rule mining always computed with the traditional consumption of transaction records analysis. In addition, the mining approach is based on a dataset of all transaction records for analysis, and then the results can be used to cross-selling or up-selling. However, this study were calculated the association rules and considered with the location factor where the products or trasaction records surrounding with sonsumers, as analyzed as the basis for the recommended product. Finally, the experimental results were also conducted the differences and compared with these two approaches.

In the experimental simulation, the transaction dataset will be divided into 12 product categories, in alphabetical A to L as a representative. Besides, there is 8 products within each product category with a number from 01 to 08 representatives, for example, the number 02 product of A product category, then namely A02 representative. If after Apriori algorithm applying, there is one rule generated to if consumer buy the 01 product of A product category, then he/she will buy the 02 product of B product category, this rule will represent as A01>B02. All the experimental dataset was divided into data sets as well as the adjacent data sets, the whole data set records for the entire simulation dataset consumption, while the adjacent dataset to simulate the consumer is located among the A, B, G, H product category area, and data collection was A, B, G, H related

consumer product category records. Finally, the experimental results are as shown in Table 2.

Rules	Entire dataset	Adjacent dataset
Data Size	394	362
Conf.>0.5	Conf.	Conf.
K03>K04	0.72	X
K04>K03	0.87	X
H01>H05	0.59	0.59
H05>H01	0.61	0.61
D07>D08	0.52	X
D08>D07	0.52	X
B05>B01	0.66	0.66
H01>H02	0.51	0.51
H02>H01	0.56	0.56
H02>H05	0.53	0.53
H05>H02	0.5	0.5
D08>D06	0.67	X
A02>B01	0.57	0.57
G01>G07	0.51	0.51
D07>D06	0.59	X
Else rules	X	3

Table 2. The experimental results.

After filtration the nearby commodity consumption records, the experimental results showed that the products with long distance of other regions or did not nearby the cunsumers will not appear in the recommendation rules. In addition, the degree of confidence value of association rules still remained high, it indicated the proposed approach has the suitable capability for products recommendation in the physical store. Finally, there were three new generated association rules for the nearby dataset, and did not appear in the entire dataset. It represented the practice value of the proposed approach.

4.5 Operational Scenario

In the first operational scenario, as shown in Figure 7(A), the RFID reader management can setup and manage the IP address, communication port, reader name, area name for RFID readers In addition, it can also obtain the data of RFID tag by connected with RFID reader via the API, including obtaining signal reader number (UID), reading time, signal strength, and signal quality. These parameters and raw data can be used to decide the physical positioning of the consumers. Furthermore, as shown in Figure 7(B), the positioning mechanism with error calculation can be evaluated by coordinate maps loading. In the next step, the system can setup the RFID tags with tracking number, and then the system can start to track RFID tags will display the calculated coordinates location. While error calculation will be calculated by the RFID tags positioning coordinates and the actual coordinates comparison by using Euclidean distance calculation.
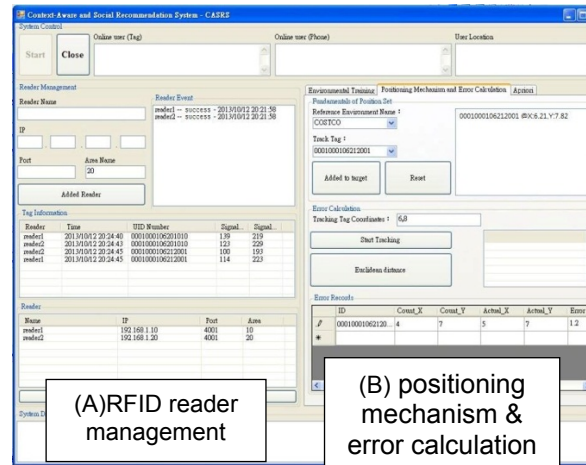

Figure 7. SRILM positioning mechanism operation screen.

In the second operational scenario, when user logins CARS via smartphone and the system will recommend products based on data registered in user's smartphone and the computing with Apriori algorithm, and then user will obtain the information on their smartphone. In this study, the system will detect the locations of users and determine different recommend products for each user with different locations through SRILM positioning mechanism. The Figure 8 shows the context information surrounding the user and Figure 9 illustrates the recommendation service on the smartphone screen.
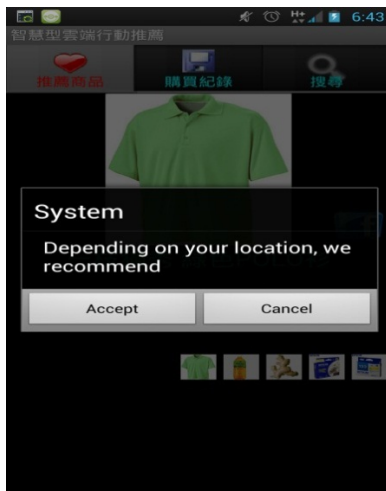
Figure 8. CARS context information.

Figure 9. CARS recommendation message.

5. Conclusion

With the concept of Software as a Service (SaaS) in cloud environment, wherever consumers are, as long as there is a platform, they can quickly use the system. Users can also use their own smartphones to operate the system. In the following, the impact and benefits on retail stores and users by using the proposed CARS system will be explored. In perspective of retail stores, product information can be sent to consumers through their smartphone, so that the value of real-time interaction can be reached and product visibility can be enhanced. In addition, store

administrator can also understand the shopping habits of consumers, and then make right marketing strategy with data mining approach. In perspective of consumers, they may decide if the featured products meet their needs based on the recommendation and their own judgment. Through smartphone, consumers can obtain real-time product information, thus increasing the convenience of shopping.

There still remains several research works need to be addressed in four issues. First, the SRILM positioning mechanism can be evaluated with other wireless techniques, such as Zigbee, Bluetooth etc. Second, the ARM can be replaced by other data mining approaches and evaluated the recommendation quality, such as decision tree and clustering algorithms. Finally, empirical research may be used to investigate user attitudes and intentions so that the system availability may be assessed.

References

[1] R. Agrawal and R. Srikan, "Fast algorithms for mining association rules," Proceedings of the 20th international conference on very large data bases, San Francisco, CA, USA,1994, pp.478-499.

[2] T. Kowatsch and W. Maass, "In-store consumer behavior: How mobile recommendation agents influence usage intentions, product purchases, and store preferences," Computers in Human Behavior, vol.26, No.4, pp.697-704, 2010.

[3] Y. L. Chen and Z.R. Chen, "A PID Positioning Controller with a Curve Fitting Model Based on RFID Technology," Journal of Applied Research and Technology, vol. 11, no. 2, pp.301-310, 2013.

[4] J. Hightower et al., "SpotON: An indoor 3D location sensing technology based on RF signal strength," UW CSE Technical Report, vol. 02, no. 02, pp1-16, 2000.

[5] W.W.T. Ngai et al., "RFID research: An academic literature review (1995–2005) and future research directions,"International Journal of Production Economics, vol.112, no.2, pp.510–520, 2008.

[6] P. Prasithsangaree et al., "On Indoor Position Location with Wireless LANs," The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication, no. 2, pp. 720-724, 2002.

[7] L.M. Ni et al., "LANDMARC: indoor location sensing using active RFID," Wireless networks, vol.10, no.6, pp.701-710, 2004.

[8] V. Alarcon-Aquino et al., "Design and Implementation of a Security Layer for RFID Systems," Journal of Applied Research and Technology, vol. 6, no.2, pp.69-83, 2008.

[9] M. S. Amin et al., "Digital Modulator and Demodulator IC for RFID Tag Employing DSSS and Barker Code," Journal of Applied Research and Technology, vol. 10, no. 6, pp.819-825, 2012.

[10] F.H. Grupe and M. M. Owrang, "Data base mining discovering new knowledge and competitive advantage", Information Systems Management, vol. 12, no.1, pp. 26-31, 1995.

[11] U. Fayyad et al., "From data mining to knowledge discovery in databases," AI magazine, vol.17, no.3, pp.37-54, 1996.

[12] P. Hadjinian et al., "Discovering data mining: from concept to implementation," vol. 1. Upper Saddle River, NJ: Prentice Hall., 1998.

[13] J. Han et al., "Data mining: concepts and techniques," Morgan kaufmann, 2001.

[14] M.Y. Chen et al., "Using RSSI simple localization method to implement the context-aware and social recommendation system," 2013 international applied science and precision engineering conference, 2013, Sun Moon Lake, Taiwan.