Architecture for Integrating Product Information and User Review on Web 2.0

Sungmoon Bae, Kwan Hee Han, Sang Hyun Choi, and Juhyun Park

Abstract—In offline markets, consumers' purchase decisions are heavily influenced by word-of-mouth. With the rapid growth of the Internet these conversations have migrated to online markets, creating active electronic communities that provide a wealth of product information. Today, customers compare the price of a product and find reviews of the product on the Internet using the smart devices. However, the information what customers want to know is not integrated and customers spend time to search proper information. This paper classifies the information into two categories – product information and user review. The product information is mainly provided by manufacturers and suppliers such as product name, specification, and price. The BOM of the product is also displayed for power users. And the detail information of parts should be provided using web services. The user review is created by other buyers who already bought and used the product. Customers share the experiences with others on the Web. User review is dynamic and can be gathered from web portals or web search engines. This paper proposes an integrated architecture for the manufacturer’s product information and the customer review on the web. It is composed of RFID system, web retrieving system and product information retriever. A prototype is also developed to demonstrate the feasibility of the proposed architecture.

Keywords—Product information, RFID, User review, Web 2.0

I. INTRODUCTION

The development of the Internet and IT technologies changes business environment and customers purchasing process. In offline markets, consumers' purchase decisions are heavily influenced by word-of-mouth. With the rapid growth of the Internet these conversations have migrated to online markets, creating active electronic communities that provide a wealth of product information. Consumers now rely on on-line product reviews, posted online by other consumers, for their purchase decisions[7][9]. Research has shown that product reviews on the Internet not only support consumers when shopping, but also lead to increased sales for retailers[18]. Radio Frequency Identification (RFID) is the latest and most advanced technology for automatic identification of objects and the most practical technology for data collection. RFID is an automatic identification system similar to barcode technology, but which acts by proximity without direct contact[1]. RFID technology, known as infrastructure of ubiquitous systems, is used to provide useful information about a product to potential customers. If a customer chooses a product what he or she wants, the pertinent information will be provided in real time. However, the information provided retailers are not sufficient to customers. They want to know more information in a detached way. They need complete information of a product.

The goal of customer satisfaction studies is to discover opinions about a company’s products, features, services, and businesses [17]. The Web contains a wealth of opinions about products, politicians, and more, which are expressed in newsgroup posts, review sites, and elsewhere [16]. Traditional purchase behaviors acquire the information that only provided by manufacturers or shopping malls. Therefore, some customers used to be able to make wrong decision because of incomplete information. To prevent those kinds of decisions, customers have to decide candidate for products beforehand what they want and then search the relevant information on the Internet. Online customer reviews constitute new channels of information acquisition. There are several approaches available which can help people make comparisons on the Web. For example, some newly emerged Web sites began to provide comparison shopping services. Shopping.com and Froogle have integrated product comparison services to provide comparative information such as price and customer reviews [21]. Users have to confirm the results of the searching for reviews such as ‘Kin-Search’, ‘Shopping Search’. With the introduction of Web 2.0 paradigm the importance of web search (including both information retrieval and information extraction) grows even more[4][23]. After that, they evaluate a price, quality and design while looking at the product. But in this case, it must still be providing incomplete information because users can’t search all relevant products or remember all relevant reviews. This paper discusses an RFID technology to resolve this problem. We suggest a framework to provide manufacturer’s product information and internet user’s reviews.

Power users who are very familiar with technology and
interest in details of a product want to know detail information of a product such as part information, when they buy some products. However, most retailers do not provide detail part information to customers. This paper also proposes an information chain for product-part relationship that is based on web services.

This paper is organized as follows: Section 2 describes related works – user review, web services, RFID, and Open API. Section 3 presents a unified framework for complete information. Section 4 demonstrates a prototype system called “Smart #” which is implemented based on the proposed framework. Finally conclusions are presented and some open issues are discussed in Section 5.

II. RELATED WORKS

The paradigm shift to Web 2.0 represents the change of the Internet from a market for goods and services to a socially centered and user-driven marketplace[22][26]. Extracting user reviews of a product becomes very important than ever. Opinion mining research to analyze opinion data on the web has become a popular topic recently[8][12]. The followings are some major researches about that subject.

Kim and et al. have proposed opinion mining method for product reviews[12]. They first do POS tagging on each review sentence, and extract feature and opinion words in form of transaction data. Then they discover association rules of needed type from the transaction data, and provide information that is summarized advantages and disadvantages using PMI-IR algorithm.

For a popular product, the number of reviews can be in hundreds or even thousands. This makes it difficult for a potential customer to read them to make an informed decision on whether to purchase the product. Hu and Liu aim to mine and to summarize all the customer reviews of a product[8]. The summarization task is different from traditional text summarization because they only mine the features of the product on which the customers have expressed their opinions and whether the opinions are positive or negative.

Some popular product review sites such as Epinions allow users to establish a trust network among themselves, indicating who they trust in providing product reviews and ratings[27]. Yeung et al. discover that users who trust each other tend to have smaller differences in their ratings as time passes, giving support to the theories of homophily and social influence. However, they also discover that this does not hold true across all trusted users. And they propose a method to estimate the strengths of trust relations so as to estimate the true influence among the trusted users.

Lim et al. provides an ranking and supervised methods that are effective in discovering product review spammers[13]. They seek to model the following behaviors. First, spammers may target specific products or product groups in order to maximize their impact. Second, they tend to deviate from the other reviewers in their ratings of products. Lim et al. propose scoring methods to measure the degree of spam for each reviewer and apply them on an Amazon review dataset.

However, less research was performed how to easily provide the user reviews to customer. This paper addresses the issue by integrating product information and user review.

Web services are currently the most promising SOA based technology. They use the Internet as the communication medium and open Internet-based standards, including Simple Access Protocol (SOAP) for transmitting data, the Web Services Description Language (WSDL) for defining services, and the Business Process Execution for Web Services (BPEL4WS) for orchestrating services[10]. Fig.1 shows the architecture of web services that contains service provider, service consumer, and service registry. Web service provider makes a web service that provides a specific function, and publishes it to public or private registry. Web service consumer finds a web service what they need in web service registry, then binds the web service and uses the service. XML and Web Services provide basic data interoperability today. Web Services, as a distributed application technology, simplifies interoperability between heterogeneous distributed systems. [20]. Liu and Gao proposed knowledge extracting platform based on web services that can solve efficiently some problem coming from nodes having data resources but lack of computing power, or some node with computing power but no data resource and low transmission bandwidth[14]. Generally, the EAI architecture pattern contains two types: Point-to-Point integration and Adapter integration[19]. Web services are applied to enterprise application integration to efficiently implement the integration of diverse applications and systems within or between the enterprises[11]. The EAIWS using Web Services can significantly change the traditional Point-to-Point approach. Using Web Services that loosely integrate applications, an application system achieves just a subsection of EAI.

![Fig.1 web services architecture](image)

RFID is a very compact technology. About as large as a pinhead, RFID tags (or simply RF tags) consist of two main components: an antenna and a chip that contains an electronic product code [17]. Radio frequency identification technology has moved from obscurity into mainstream applications that help speed the handling of manufactured goods and materials. RFID enables identification from a distance, and unlike earlier bar-code technology. RFID tags support a larger set of unique identifiers than bar codes and can incorporate additional data such as manufacturer, product type, and even measure environmental...
factors such as temperature [25]. RFID reader uses to extract data from a register in an RFID tag can also be applied to collecting sensor-derived data [24]. High-end fashion retailers are beginning to weigh how RFID-based systems might improve the overall customer-service and consumer shopping and spending experience [16]. Chen et al. used RFID technologies on the food produce traceability system. Using RFID technology will be easy to trace each object, not only for the goods lots. RFID technology can also record all events automatically and acquire the information about the food production by handhold devices [3].

A variety of application system is appearing using the RFID technology. Especially, distributors are introducing the Smart-shelf make it possible to check the stock information and product information in real time, have increased. Smart-shelf that means the system displays the product information using the RFID technology. When put a product up on the Smart-shelf, the product information is quick and easy acquired to the screen. Real-time information can be obtained through smart shelves, which have inbuilt RFID tag receivers. This would help retailers track the exact number of products they hold [15].

Open API that is Web 2.0 technology to open own corporate API in public. Wikipedia defined as API as “An application programming interface is a source code interface that a computer system or program library provides in order to support requests for services to be made of it by a computer program.” And Open API is to open that a variety of users can use it through the web and it is so characteristic of open oriented. Generally, it is in use mixed similar implications as Open service API, Web services API (Google APIs; Daum Open API).

III. ARCHITECTURE FOR INTEGRATING PRODUCT INFORMATION AND USER REVIEW

A. User Reviews and Product Information

When customers visit offline stores, they want to buy a right product. They, however, are only provided restricted product information by sales person. The information provided by a sales person could be insufficient to make right decision. They need more information that is not provided by a sales person. The need complete information of a product. Complete information of a product that is composed of supplier-side information and customer-side information. Supplier-side information, provided by manufacturers or retailers, is a product specification, price, etc. Customer-side information, provided by existing user who has used the product, is customer reviews, user rating, comments, etc.

Complete product information is composed of following two categories. 1) Product information: includes product features, specifications, BOM structure, and price of the product. The information is provided by manufacturers or retailers. 2) Review information: includes reviews of existing users, ratings of the product, and comments. The information is created by internet users and communities and extracted by web search engines.

B. Integrated Architecture

Fig. 2 shows the integration architecture for complete product information. The architecture is composed of four parts – user interface, application logic unit, product information gathering, and customer review gathering. Recent approaches successfully use smart phones to directly relate products (e.g. via barcode or RFID) to corresponding reviews, making these available to consumers on the go[20]. The user interface gets a product id using RFID attached to a product and hands over to application logic unit. Customers just go through the product and the RFID base input system gets the product id from a tag on a product. The application logic unit receives a product id, and requests product information and review information using the product id. The product information is retrieved from product database in the store. If users want to know detail BOM structure of a product, the application logic unit gathers the part information by navigating the supply chain of a product. Because legacy database systems are developed in various platforms, it is needed that a standardized method to access legacy systems. The review information is extracted from a web search engine using open APIs. The information is managed by internet communities in web portals. A search engine only retrieves the information and passes the result to the application logic unit. The application logic unit binds the product information and review information, and sends them to user interface. Finally, the user interface displays complete information of given product.

C. Information Chain of Product-Part Relationship

A product is composed of many parts that are provided by suppliers on a supply chain. A product has its own BOM structure that contains a product-part relationship. Each company on the supply chain manages some part information – part name, lot number, quality information, and so on. However, their system are not homogeneous but heterogeneous. They use
different hardware, operating system, database, and programming language.

In this paper, we propose an integration framework for supplier’s heterogeneous systems. Web services provide interoperability in heterogeneous computing environment by XML based standards. They wrap the legacy application with WSDL and SOAP standard. Each supplier should provides XML-based web services which get part id as an input and return part information as result. To get the sub-part information, it also calls a web service provided by sub-part suppliers. We call this framework as information chain of product-part relationship. The framework is shown in Fig 3.

D. Open API of Web Search Engines

Information extraction (IE) business through the web is defined by its input and its extraction target [16]. Its input usually is represented in free text and its output is represented in standardized formats such as XML. Input and extraction target of Open API is also entered and got a result by these formats and data search can be automated under this way. In particular, if a customer selects the category and attribute predefined to acquire some results of the specific product, the customer can acquire its result in real time.

When a customer will decide into purchasing online through the web, it will be able to make his or her comparisons such as specification, price of a me-too product. If offline shopping mall adopts this comparison shopping, a customer can check online review while looking at the real product on display. When a customer pick up a product on the antenna, can see review information of it. Accordingly, automated search through web ensure high level of satisfaction. It can shorten the decision-making time and the verification of incomplete information is possible. Currently, Google and Amazon provide Open API to review of product. This paper introduces the example that uses Google APIs.

IV. IMPLEMENTATION OF PROTOTYPE

A prototype of the system named ‘Smart #’ that provides the complete information implemented using the Smart-shelf, Open API. When the antenna of Smart # detects a product, it shows the product information in the product DB and the review information through the Open API to the client in real time. Fig. 4 shows the system architecture of the prototype. Smart # has been developed the following development environment.

- Application Server: Windows Server2003
- Client OS: Windows XP
- Development tool: Microsoft Visual Studio .NET 2005
- Web server: Microsoft IIS V5.1
- OpenAPI: Daum OpenAPI
- DBMS: Microsoft SQL Server 2000
- RFID development kit: SmartNL-RF300

Smart # is composed of the following components.

- Product: The RFID tag is embedded in the product which has unique ID.
- Antenna: When a product is detected within a certain distance from the antenna, server gets the corresponding product ID through the reader and client.
- Reader: This is a machine that reads the information of tag and information that collected by reader is transferred to the middleware.
- Client: This gets a product ID through the reader and transfer the product ID to server.
- Application Server: This reads product information that corresponds to the product ID that received from the client from the product DB and reads the review information corresponds to the product through Open API. Then these complete information transfer to the client.
- Product DB: This stores product information such as video advertising, product image, price, manufacturer’s name, weight of product.
- Open API: We use Daum APIs that provides review information.

Fig. 5 RFID kit: smartNL-RF300
A. RFID kit for Detecting a Product

RFID Kit used for Smart# has applied SmartNL-RF300 model in the band of 900MHz (Fig. 5). It is highly portable as a reader-antenna integrated USB type and is easy to be connected to PC and shows high reading ability within the range of 20~30cm.

B. Data Model

Smart# Database table structure design of the system, as shown in Fig. 6 and based on a relational database and all databases can be easily applied. There are Company Table where shows the information of the product manufacturer and Product Table where shows the product details and BOM Table where shows the information on the components constituting the product. As the components in BOM Table have to be expressed in a tree structure by establishing a link between subordinates and superiors, there is Relation Table to define the relations in each component. In this paper, we design for database tables for TV, PC, Computer, and enter the same virtual products and BOM data to be displayed on the Smart# OpenAPI was extracted from a review of information.

For Product Table, there are data such as the number, name, price, volume, manufactured date and features of the model and RFID tag numbers so as to allow searching the product details for the relevant tag by tag recognition. In addition, the constituent parts of its products to list data can be found in the BOM table. BOM table, all the components of products such as parts specifications, manufacturing companies, materials, features has been entered.

With Relation Table, BOM relations of the product recognized by parent component codes and child component codes, which define the assembly rule for each BOM, can be presented in a tree structure. BOM tree structure in several levels can be expressed in a simple table structure as it has been modeled into a circular structure. Fig. 7 shows BOM structure in a tree view, images and detail information of chosen component.

C. Sequence Diagram of Prototype

Fig.8 shows a sequence diagram of the system. The objects that constitute the system can be roughly categorized into Client, Smart#, Application Server, Product DB and Open API and events occur between objects with the downward course of time.

When the product tag gets closer to Smart# by Client, it will be read by RFID antenna and the tag information can be verified. Application Server collects product data and BOM data through searching the product information from Product DB with tag number after verifying the effectiveness of recognized tag and displays it on Smart# screen. If it is unregistered or unavailable tag, it indicates that it is unregistered tag onto the screen. When it requests to Open API in http format with the product name.
searched from DB, data for relevant product receives a response in XML format and Application Server indicates the information on product comments for each shopping mall onto Smart# screen. Therefore users can verify product details provided by seller, BOM information provided by manufacturer and product comments provided by buyer all at once from Smart# screen through product tag recognition.

D. Daum OpenAPI

Daum OpenAPI(http://apis.daum.net) opens various data platform and contents such as shopping, maps, searching, RSS, cafes, blogs and so on to the public so diverse service can be provided from this. Product details and user’s comments were obtained from implemented Smart# with Daum Shopping API.

Table1. XML response element

<table>
<thead>
<tr>
<th>Title</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>description</td>
<td>description of result</td>
</tr>
<tr>
<td>totalcount</td>
<td>total number of result set</td>
</tr>
<tr>
<td>sort</td>
<td>sort order of result</td>
</tr>
<tr>
<td>result</td>
<td>the number of results displayed in a page</td>
</tr>
<tr>
<td>target</td>
<td>target of search</td>
</tr>
<tr>
<td>pageno</td>
<td>page number of result</td>
</tr>
<tr>
<td>q</td>
<td>search word</td>
</tr>
<tr>
<td>item</td>
<td>information of the individual search results</td>
</tr>
<tr>
<td>docid</td>
<td>ID of product</td>
</tr>
<tr>
<td>title</td>
<td>trad name</td>
</tr>
<tr>
<td>description</td>
<td>Description of product</td>
</tr>
<tr>
<td>price_group</td>
<td>product information of the group</td>
</tr>
<tr>
<td>price_min</td>
<td>lowest amount</td>
</tr>
<tr>
<td>price_max</td>
<td>best amount</td>
</tr>
<tr>
<td>publish_date</td>
<td>release date</td>
</tr>
<tr>
<td>maker</td>
<td>maker</td>
</tr>
<tr>
<td>brand</td>
<td>band</td>
</tr>
<tr>
<td>image_url</td>
<td>image path</td>
</tr>
<tr>
<td>shoppingmall</td>
<td>the number of shoppingmall which contains reviews</td>
</tr>
<tr>
<td>review_count</td>
<td>the number of reviews</td>
</tr>
<tr>
<td>link</td>
<td>url of the product</td>
</tr>
</tbody>
</table>

If you request to the address such as http://apis.daum.net/shopping/detail with the product name searched from DB on Smart#, you get a response in XML format. The request variable should be delivered at this point, apikey value and product name are available for the request variable. apikey value is the invested eigen key value so as to use Open API and the product name is the name value of the product searched from tag. Smart# can display the product details in a list format by interpreting received XML. Table.1 shows XML elements received from Open API and it includes various product data such as the bottom price/ceiling price, specifications, number of reviews, shopping malls, launched date and so on. Fig. 8 presents C# code to be received in XML format after requesting through Open API. It requests Open API the product name searched from DB by tag number through URL with invested apikey value and the encoded value. After the request, XMLElement object will be returned and product details and review information will be obtained by collecting information of various item elements affiliated to channel elements in foreach statement. Fig. 9 is an example of TV product which has been presented in XML format, and you can see that the data of TV product has been presented in the same elements as docid, title, price_min, price_max, publish_date and so on which are subordinate to item elements.

![Fig. 9 XML form of review example](image)

If you request to the address such as http://apis.daum.net/shopping/detail with the product name searched from DB on Smart#, you get a response in XML format. The request variable should be delivered at this point, apikey value and product name are available for the request variable. apikey value is the invested eigen key value so as to use Open API and the product name is the name value of the product searched from tag. Smart# can display the product details in a list format by interpreting received XML. Table.1 shows XML elements received from Open API and it includes various product data such as the bottom price/ceiling price, specifications, number of reviews, shopping malls, launched date and so on. Fig. 8 presents C# code to be received in XML format after requesting through Open API. It requests Open API the product name searched from DB by tag number through URL with invested apikey value and the encoded value. After the request, XMLElement object will be returned and product details and review information will be obtained by collecting information of various item elements affiliated to channel elements in foreach statement. Fig. 9 is an example of TV product which has been presented in XML format, and you can see that the data of TV product has been presented in the same elements as docid, title, price_min, price_max, publish_date and so on which are subordinate to item elements.

![Fig. 9 XML form of review example](image)

E. User Interface

The User Interface was implemented in the Visual C# Form. Main form is consisted of a tap that sets RFID and a tap that displays the product details. RFID setting tap connects the middleware to RFID reader and can operate and control the antenna and display the details of recognized product on a list. Fig. 10 shows the screen on which the product details of implemented Smart# system have been indicated. It shows empty screen before recognizing the tag, then displays the product details automatically when it recognizes the product tag.

![Fig. 10 screen shot of prototype](image)
Tag will not be recognized by the reader while the product details are being displayed. The screens that display the product details can be generally categorized into three types that are as follows:

- **Product Information:** information is provided by the seller. It displays information such as trade name, product images, price, manufacturer, manufacturing date, production characteristics. When you select the manufacturer name, company homepage will appear.

- **BOM Information:** Provided by the manufacturer that the product is about the constituent parts. Represented as a tree structure format, the configuration of the parts can be easily identified.

- **Review Information:** It displays a list of shopping malls where sell relevant products along with the bottom/ceiling price and number of reviews for the products. By clicking a shopping mall with reviews, user can move to the relevant shopping mall so he can find more details about the products and a list of buyer’s reviews as well. Fig. 11 indicates a list of buyer’s comments on the product and shows its ratings along with the comments on the product.

```csharp
XmlDocument doc = new XmlDocument();
XmlNode node, n;
string url = "http://apis.daum.net/shopping/search?apikey=00000000000000000000000000000000QWERTYUIOP"
.doc.Load(string.Format(url + @"/node.SelectSingleNode("rss");
node = doc.SelectSingleNode("channel");
foreach (XmlNode el in n.SelectNodes("item"))
{ ListViewItem li = new ListViewItem();
 string pprice= el.SelectSingleNode("title").InnerText;
 }

Fig. 11 request OpenAPI and XML data extraction

V. CONCLUSION

User reviews on the Internet not only support consumers when shopping, but also lead to increased sales for retailers. However, this information is scattered and needs some effort to retrieve. In this paper, we have defined complete product information of a product that is composed of supplier-side information and customer-side information on the web. The supplier-side information contains features, specifications, a price, and a BOM structure. The customer-side information contains the user reviews. Part suppliers provide web services for retrieving part information. Users can navigate a BOM structure and gather some part information using web services. We proposed integration architecture for complete product information. The proposed architecture is based on the RFID-enabled mobile PDA. OpenAPI of web search engine and web services. RFID technology requires minimum user interaction and provides easy way to get product id. Web services provide interoperability of heterogeneous systems in supplier chain. In that case, part suppliers play a role of web service provider. Web search engines such as google, naver, daum and so on gather user reviews in the Internet. The architecture enables customers to do right decision for buying a product. They can access all information needed using only one device – mobile PDA – without any inconveniences.

REFERENCES


