Mobile Learning System Using Multi-dimension Data Warehouse Concept-Based on Botanical Data

Rong-Jyue Fang[1], Howard Lo[2], Yu-Chen Weng[3], Hua Lin Tsai [4]
[1]Chair Professor, Department of Information Management, Southern Taiwan University of Technology, Taiwan
[2,4]Graduate student, Department of Industrial Technology Education, National Kaohsiung Normal University, Taiwan
[3]Teacher, Linyuan Elementary School, Kaohsiung county, Taiwan
13F., No.225, Guanghua 3rd Rd., Kaohsiung City 806 Taiwan

Abstract: - The development of mobile wireless technology had generated a significant effectiveness among practitioners and academics because it results in shifting the pedagogical environment from traditional settings to mobile learning settings. In the beginning of 21st century, mobile learning will become the next step of evolution of ICT(Internet and Communication Technology) pedagogical methodology and tools. New ICT technology will be embedded in every corner of society or pedagogical setting. Botanical learning activities are ubiquitous in the subject of Science and Technology of the Grade 1-9 Curriculum, in Taiwan. While most of the teachers were suffered from a lack of botanical knowledge, we applied the concept from data warehouse and document warehouse to construct a MDBDIS system to fulfill their information and pedagogical need. In the result of this study has indicated how mobile learning technology can be integrated into botanical data warehouse, the recall and precision rates identified the efficiency of the MDBDIS system.

Key-Words: - Mobile Learning, Data Warehouse, Document Warehouse, Botanical Course

1 Introduction

In the beginning of 21st century, M-learning (Mobile Learning) will become the next step of evolution of ICT Pedagogical methodology and tools. Many types of portable devices have been applied in the field of learning and education. New ICT technology will be embedded in every corner of society or pedagogical setting.

Portable devices include a wide range of information technologies such as mobile phones with Internet connection to Personal Digital Assistants(PDAs), and handheld computers designed to be personal information managers[3] . They are highly portable and personal computing appliance, which can be carried around and used 'any time, anywhere'. They can be used for a variety of functions; for example, to manage work or study schedules, to record and store data, and to access and disseminate information[4] .

Several studies have been conducted to probe the potential use of portable device as learning tools in pedagogical settings[5] [6] . There is great potential for portable devices to provide students with a tool that can support learning in various contexts[4] . They can also provide access to learning resources anytime and anywhere.

The benefits of the M-learning have extended the learning environment and integrate it with real life environments, where learning can occur in a campus or educational settings. Our study tries to adopt the advantage of portable devices into botanical courses of elementary settings. Since the preparation of tomorrow's citizens is incumbent on the elementary schools, which ought to include the new technological applications in the educational process[7] . We integrate M-learning into botanical courses based on the view of naïve users' view.

Botanical learning activities are ubiquitous in the subject of Science and Technology of the Grade 1-9 Curriculum, in Taiwan. We have constructed lots of learning resources in our school website including glossaries of botanical terms and pictures classified by the species, genus, and family of the plants or by order of Chinese character strokes of the plants. There are also omnipresent botanical resources dispersed on the Internet. But Lo's[8] research found

it was difficult for naïve user to search related information from the Internet, because there is too much information on the Internet, and it is too much to read, or even skim. There is a need for effective tools that allow user to search the information they desired whether on the Internet or a database server.

In this study, we applied Data Warehousing[9] concept to develop a multi-dimensions botanical documents inquiry system (MDBDIS) to serve naïve users' information need. Different from the conventional plant documents inquiry system based on the classified glossary of Botanical Terms, such as species, genus, family, order, class, phylum, and kingdom, we adopted the intuitive feature of plants such as Phyllotaxy, Leaf shape, Apex, Base, Margin, Inflorescence as the input of the MDBDIS system. Users can retrieve the information they needed from the database via the query based on their observation on the feature of the plants.

2 Mobile Devices for Tutoring and Individualized Learning Processes

Mobile devices like 3rd Generation (3G) cellular phones and PDAs bring new dimensions in learning and education. M-learning extends and integrates to the real pedagogical environment. What is the advantage and disadvantage of m-learning? How can m-learning integrate into conventional pedagogical contexts.

2.1 The advantage of the mobile device

The benefits of the newly gained mobility of mobile device are expected to be reflected in more efficient education and improved learning results. Kinshuk [3] reported the potential of mobile learning as a way to change the learning activities brought by the possibility to access any of the information that is available through the Internet. When using mobile devices, students are able to construct useful knowledge in a real situation anywhere and anytime. High Portability and small size are another advantage of mobile pedagogical or learning tools.

2.2 The disadvantage of the mobile device

According to the study of Waycott et al [4] , the first weaknesses of the mobile devices are small screen size. The screen of handheld devices is too small for user to read information from it. The second limitation is that navigation on portable devices was difficult. Awkwardness of entering text is another disadvantage of mobile devices. Users found the methods for entering text on the PDA or any mobile devices to be slow and awkward. Tapping out letters on the portable devices' small onscreen keyboard was slow and error-prone.

2.3 Why m-learning is needed in the botanical pedagogical activities

Chang et al. [2] present the concept of the AD Hoc and Mobile Classroom, and eSchoolbag systems. With these systems one can construct a ubiquitous learning environment to support both indoor and outdoor activities in learning. In our outdoor botanical pedagogical activities, mobile devices can assist our teachers and students to learn the botanical knowledge anywhere and anytime.

2.4 The limitation of contemporary botanical query system

Botanical pedagogical activities are omnipresent in the subject of Science and Technology of the Grade 1-9 Curriculum, in Taiwan. But unfortunately, most of the teachers are not familiar with the botanical domain. Our teachers try to search related information from several sources, but they found awkwardly to search related information from the resource listing as follow:

1. School based botanical web page: Most schools have constructed school based botanical pedagogical web pages. But most of them are listing by the order of Chinese character strokes of plant's name, or by order of taxonomy of botanical classification. It is difficult for naïve user without any botanical knowledge to search information they needed.
2. Search Engine on the Internet: Search Engine is popular tools for user to search information they needed. But we found the same awkward situation we talked above; users don't know how to key in the keyword about the plant they observed.
3. Professional botanical inquiry system: There are several professional botanical query system have been constructed based on the taxonomy of botany, such as Database of Native Plants in Taiwan(DNPT), constructed by Herbarium, Institute of Botany, Academia Sinica. The database includes bountiful of information about native plants. Users need the knowledge of botanical taxonomy, such as Scientific Name, Family Name, Genus Name, Chinese Name, or Elevation Range as an input to inquire the related information from the DNPT system. It was impossible for naïve user to implement such a query.
4. Paper based botanical documents: The traditional ways to find an answer is to skim on

a paper based books. We can compare the image we observed with the one on the illustrations of botany, but it takes time and is difficult to find related information they need either.

2.5 The purpose of our study

The aim of our study is to apply the concept of data warehouse and document warehouse to develop a multi-dimensions botanical documents inquiry system (MDBDIS) based on the users' observation. User can exploit the intuitive feature of plants such as Phyllotaxy, Leaf shape, Apex, Base, Margin, Inflorescence as an input of the MDBDIS system. User can use their mobile device to inspect the plants anywhere, anytime in the campus and the neighborhood.

2.6 The limitation of our study

In this study, we build a wireless environment in our campus and the neighborhood. There are several limitations about our MDBDIS system:

1. The wireless environment is limit to Shanlin Elementary School, Koahsiung County, Taiwan and the neighborhood.
2. The test data were from (1)200 botanical documents from Academia Sinica, Taiwan, (2) 150 documents from Council of Agriculture, Executive Yuan, Taiwan, (3) and 150 documents from user self-constructed web pages.

3. Multi-dimensions Inquiry System for Plant Document

In this study, we apply the concept of data warehouse and document warehouse into our Multi-Dimensions Botanical Documents Inquiry System .

3.1 The concept of Data Warehouse

Data warehouse is the latest must-have marketing weapon in Business domain. It provides architectures and tools for designers to systematically organize, understand, and use their data to support their strategic decision. We found it is also useful for information retrieval.

3.1.1 What is a Data Warehouse?

The construction of data warehouse, which involves data cleaning and data integration, can be viewed as an important preprocessing step for data mining [10] . Data warehousing provides architectures and tools for decision-makers to systematically organize, understand, and use their data to make strategic decisions. Data warehouses have been defined in many ways, making it difficult to formulate a rigorous definition. According to W.H Inmon, a lead scholar in data warehouse domain, "A data warehouseis a subject-orient, integrated, time-variant, and nonvolatile collection of data in support of managements' decision making process" [11] .

3.1.2 Difference between Operational Database Systems and Data Warehouse

Since most users are familiar with relational database systems, it is not difficult to tell the difference between both systems by comparing these two kinds of systems. The major task of relational database systems is to conduct an on-line transaction and query processing, called OLTP. OLTP stored the transaction data into relational database system day by day. On the other hand, Data warehouse systems can organize and present data in various ways in order to assist the decision-maker. These systems are known as on-line analytical processing (OLAP) systems[10] . OLAP includes several multidimensional data concept. With the multi-dimensional concept, data areorganized into multiple dimensions, and each dimension contains multiple levels of abstraction defined by concept hierarchies. These hierarchies endow OLAP with useful function, such as Roll-up, Drill-down, Slice and dice. Fig. 1 depict a 3-D data cube.

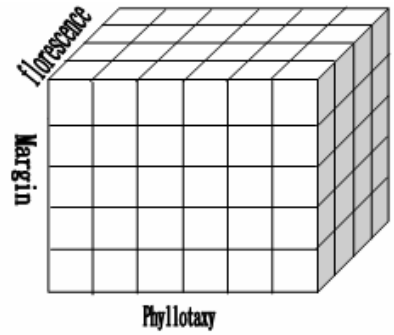

Fig.1 A 3-D data cube

Fig. 2 depict a slice view of data in the Data warehouse, and Fig. 3 show another view of dice to represent the data in the Data warehouse.

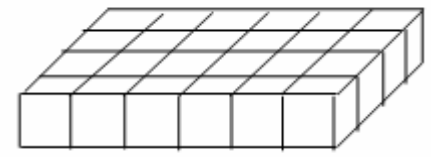

Fig. 2 slice operation

3.2 The concept of Document Warehouse

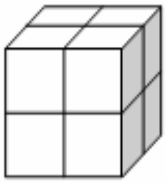

Fig. 3 dice operation

According to the report of Survey.com (http://www.survey.com), 20% of information was stored in the form of structural data, like relational database; 80% of information was came from non-structural data, such as text. As we delve into the realm of text-oriented business intelligence, we will quickly see two parallels with more traditional numeric business intelligence operations. First, in both cases we are dealing with large quantities of information, since 80 percent of all business information is textual, documents warehouses will have to deal with massive amounts of raw data [12] . Since there are bountiful botanical text resources on the Internet, we need a good tool to search and retrieve the related information from the Internet. According to Sullivan[12] , "If the web is the mine from which we extract raw material, document warehouses are the mills and refineries through which we turn text resources into usable business intelligence and distribute it to the ultimate customer." Applying the concept of document warehouse into our botanical query system is the aim of this study.

3.2.1 The architecture of the document warehouse

There are several primary component of the document warehouse [12] :

1. Document sources: Finding documents to store in a document warehouse is never a problem. We are surrounded by text. Generally, there are three types of sources such as internal sources, the Internet, and subscription services.
2. Text processing servers: within the document warehouse, there are four distinct types of servers, such as document collection servers, text analysis servers, publishing and distribution servers, and storage servers.
3. Text bases and other storage repositories: text bases, databases with support for support for text, file systems, and combined file system/database repositories are several options for storing text in a document warehouse.
4. Metadata repositories: metadata is information describing documents and texts and is a critical piece of the document warehousing environment. It serves the following purposes:(1) Improves search precision and recall, (2) Allows for extended searching options, such as by author, date of publication, and so on. (3) Categorizes texts, (4) Indicates relative levels of quality, reliability, and timeliness.
5. User profiling: the document warehouse allow user to specify that particular resources be monitored via user profiles. There are three methods for describing the users' interests, such as by subject, by source, or by keyword.

3.2.2 The basic step to construct a data warehouse and document warehous

As Fig. 4 shows, the basic step for building data warehouse and document warehouse are similar. The

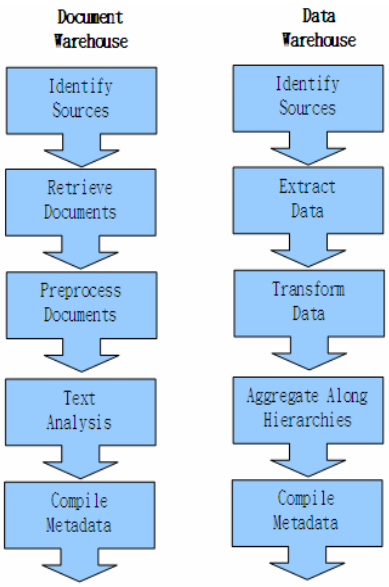

Fig. 4 The basic steps in data/document warehouse [12]

five main steps are

1. Identifying sources, the only difference is the type of source. Structured data is the sources of data warehouse; and non-structured data, like text documents will be the source of document warehouse.
2. Retrieving information from the warehouse. Documents will be the result for document warehouse; and summary, structured data will serve as the result for data warehouse.
3. Preprocessing stage is to transform data for data warehouse, or to format the information in a consistent manner to support later operation in the document warehouse process.
4. Text analysis is the heart of Document warehouse, and data mining is the analysis step for Data warehouse.

5. Compiling metadata is a critical operation because metadata makes explicit, in easily queried form for information retrieval.

3.3The multi-dimensions botanical documents inquiry system

In this study, we apply the concept derive from data warehouse and document warehouse to construct a multi-dimensions botanical documents inquiry system to assist the naïve users. Users can observe the features of a plant such as Phyllotaxy, Leaf shape, Apex, Base, Margin, Inflorescence as inputs of our Multi-Dimensions Botanical Documents system. With the mobile device like 3G mobile phone or PDA, users can easily found the related information about the plant.

3.3.1 The methodology of MDBDIS

Similar to the concept of data and document house, we need to process the original documents first. The next step is to encode the document. The third step is to build a metadat in order to retrieve the document in the document warehouse. On the users' side, they can use they 3G mobile phone, PDA or a laptop with wireless device to search information from MDBDIS. Fig. 5 show the model of MDBDIS.

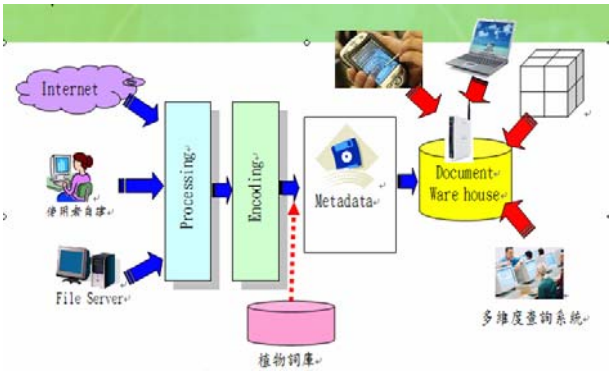

Fig. 5 Model of MDBDIS

3.3.2 The algorithm of MDBDIS

We exploit the relational algebra[1] as the algorithm of MDBDIS. It's very common in database domain. The example of the simple query algebra is depicted in Fig. 6. If user want to find information related to some kind of plant with "serrate" feature. The algebra will be like equation (1).The SQL like algorithm will be Equation(2).

$$\sigma_p(R) \quad \sigma: \text{Select}, \quad {}_p: \text{Predicate}, \quad R: \text{Relational table}$$

Fig. 6 relational algebra

$$\sigma\, \text{LeaveMargin} = \text{serrate}(\text{Plants}) \quad \text{Equation (1)}$$

$$\text{s where Plants}(\text{name}: \text{s}, \text{LeaveMargin}: \text{"serrate"}) \quad \text{Equation(2)}$$

4 Efficiency and effectiveness

When considering query performance evaluation, we should first consider the inquiry task that is to be evaluated. Efficiency and effectiveness is the most important part of evaluation work. We adopted the recall rate and precision rate to assess MDBDIS system[1] .

4.1 Recall rate

Recall rate is to caculate the fraction of the relevant botanical documents which has been retrieved. The Equation (3) depicted the caculation of recall rate. In our MDBDIS system, The recall rate of LeaveMargin = "Serrate" and florescence="pink" would be 0.83, for 5 relevant documents had been retrieved, and there 6 relevant documents in document warehouse.

$$\text{Recall} = \frac{|\text{number of relevant Docs retrieved}|}{|\text{number of relevant Docs in collection}|} \quad \text{Equation (3)}$$

4.2 Precision rate

Precision rate is the fraction of the retrieved botanical documents which is relevant. The equation 4 will indicate the precision rate. The precsion rate of LeaveMargin = "palmately lobed" and Phyllotaxy =" opposite" is 075. There are 3 relevant documents had been retrieved and 4 documents had been retrieved.

$$\text{Precision} = \frac{|\text{number of relevant Docs retrieved}|}{|\text{number of Docs retrieved}|} \quad \text{Equation(4)}$$

5 Conclusion

In this study, We integrated mobile learning technology into bbotanical pedagogical activities. Since most of the teachers were suffered from a lack of botanical knowledge, we applied the concept of data warehouse and document warehouse to construct an MDBDIS system to support the m-learning pedagogical activities.

This paper has indicated how mobile learning technology can be integrated into botanical

pedagogical activities, the recall and precision rates showed the efficiency of the MDBDIS system. Via MDBDIS, the naïve users can easily key in the features of a plant, such as Phyllotaxy, Apex, Base, and LeafMargin, as a input of the query, then the system will narrow down the number of relevant documents to meet user's information need.

In this paper, our finding have provided evidence of an empirically applied mobile learning technology into botanical data warehouse that is a possible new avenue for future study. In future research, we want to try to integrate artificial intelligence into MDBDIS system to boost the efficiency and effectiveness of mobile learning. We also want to extend our research to the other domain such as entomology, zoology.

References:

[1] Baeza-Yates, R. & Riberio-Neto, B. *Modern Information Retrieval*, Addison Wesley, 1999

[2] Chang, C. Y., Sheu, J. P. & Chan, T. W., Concept and design of Ad Hoc and Mobile classrooms. *Journal of Computer Assisted Learning* V.19(3), 2003, pp. 336.

[3] Kinshuk, Suhonen,J. Sutinen,E., & Goh, T., Mobile Technologies in Support of Distance Learning, *The Asian Society of Open and Distance Education,* Asian J D E, Vol. 1, No. 1, 2003, pp. 60-68.

[4] Waycoot, J. & Kukulska, H.A, Student's Experience with PDAs for Reading Course Materials, *Personal and Ubiquitous Computing*, Springer London, 2003, pp. 30-43.

[5] Hennessy S., Graphing investigations using portable(palmtop) technology, *J Computer Assisted Learning*, 2000, pp. 243-258.

[6] Robertson, S., Calder, J., Jones, A., & O'Shea, T., The use and effetiveness of palmtop computers in education, *British J Educational Technology* , 1997, pp. 177-189.

[7] Panagiotes S., & Anastasiades, Distance learning in elementary schools in Cyprus: the evaluation methodology and results, *Computers & Education*, Vol.40, No.1, 2003, pp.17-40.

[8] Lo, H., Applying Artificial Neuro Network in Plant Documents Clustering, *Paper presented at 2005 International conference of Curriculum & Instruction in Technology Education,* Taiwan. 2005.

[9] Inmon, W.H., Rudin, K. Buss, C.K., & Sousa, R., *Data Warehouse Performance. John*, Wiley & Sons, Canada, 1999, pp.4-10.

[10] Han, J. & Kamber, M. , *Data Mining : Concetps and Techniques*, 2001, pp. 39-41.

[11] Inmon, W.H., *Building the Data Warehouse*, John Wiley & Sons, New York, 1996.

[12] Sullivan, D., *Document Warehousing and text mining*, John Wiley & Sons, 2001, pp.56-58.