Communications of the IIMA

Volume 5 | Issue 3 Article 6

2005

Mining User's Preference Information through System Log toward a Personalized ERP System

Zhang Qun
School of Management, University of Science and Technology

ShouLin Lai
School of Management, University of Science and Technology

ShuFen Dai
School of Management, University of Science and Technology

Xuedong Gao
School of Management, University of Science and Technology

TieKe Li
School of Management, University of Science and Technology

Mining User's Preference Information through System Log toward a Personalized ERP System

Zhang Qun
School of Management, University of Science and Technology Beijing,Beijing,100083
Tel +86-10-62332744 Fax +86-10-62333582 Email: zq@manage.ustb.edu.cn

ShouLin Lai
School of Management, University of Science and Technology Beijing,Beijing,100083

ShuFen Dai
School of Management, University of Science and Technology Beijing,Beijing,100083

XueDong Gao
School of Management, University of Science and Technology Beijing,Beijing,100083

TieKe Li
School of Management, University of Science and Technology Beijing,Beijing,100083

ABSTRACT

This paper discusses how to mine the user's preference information through the system log of the ERP system. Our findings support that the user's preference degree is positively correlated with the usage frequency and negatively correlated with the latest waiting time. With a derived function of user's preference degree, it can be applied to guide the automatic readjustments to the system interface and content accordingly. Taking a contract management sub-system as an example, the paper shows how to mine the system log toward a intelligent personalized ERP system.

Key words: system personalization user's preference degree data mining system log.

INTRODUCTION

ERP system is essentially a system of information service for individual users in the enterprise. The individual user's degree of satisfaction, therefore, is an important criterion to measure the success of any ERP system. How to meet the uncertain personal needs has become a challenge for developers of various ERP systems as well as a key role that the ERP system will play in the future. With the development of service trades, the customer's personal demands have become increasingly diversified, which in turn makes it more difficult for software developers to design similar systems for many different individual users.

This paper attempts to show how to mine the users' preference information through the system log, then this preference degree information can be applied to guide the automatic readjustments to the system interface and content. The system log is the only data source to mine user's personal demands.

THE STRATEGY OF SYSTEM LOG MINING

In order for the system to mine the user's preference information in an automatic way, the system log is the only data source. By defining the structure of the system log, the user's preference structure can be mined according to their access log. And then their preference structure can be applied to adjust/modify the style of interface and content of the system, so as to meet the uncertain demands of most customers intelligently.

The system log mining is mainly divided into log content mining, log structure mining, log usage mining (Xinzhong Chen, et. al, 2002), as shown in Figure 1:

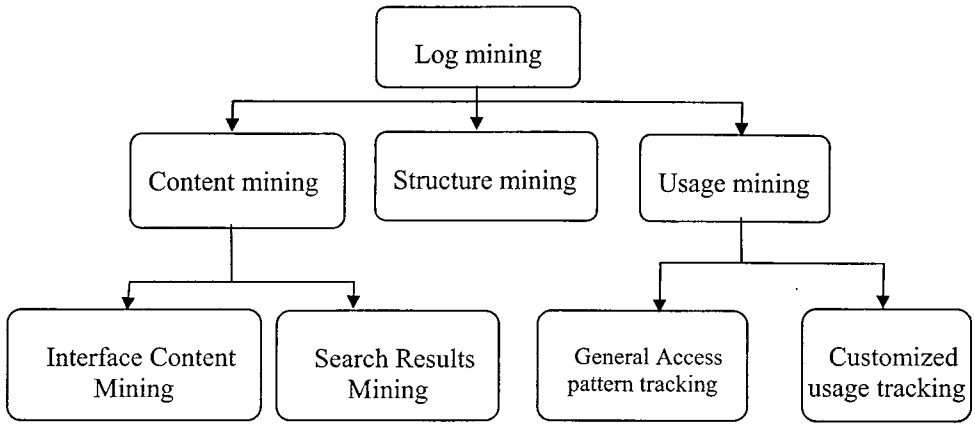

Figure 1: The mining of system log.

The examples described in this paper are log usage mining, which usually falls into two categories: the tracking of general access pattern and the tracking of customized usage. With the tracking of general access pattern, the system analyses the users' access pattern and tendency through the log so as to give the better structure and grouping of resource providers. For example, through the analysis of system log, the system obtained the information that the employee in charge of chemical products has the tendency of packaging with 500ml bottles in "how to package" item of the contract, and omitting the producer's name on the printout of the contract, and then these will enter into the knowledge rule base. Another example would be that once the system obtained the information that the employee in charge of the electrical equipment usually selects 'to package with crates' and usually print out the producer's name on the printout of the contract, and then these will enter the knowledge rule base as well.

With the customized usage tracking, the system analyses and tracks the preference of an individual user by mining the system log and to design the personal style of appearance and colour for the individual user on the basis of the tracking pattern.

The general logical processing consequence of log mining is exhibited in Figure 2: the preliminary processing of the original log file, elimination of useless data, screening out the related items generated by the customers, mining the related items and deriving out the knowledge rules. The knowledge rules will be applied to the real-time recommended engine; once the user applies the system again, the latest and most probable content items obtained will be shown. As soon as the user finishes the session, the new related items will enter the log table as the new record.

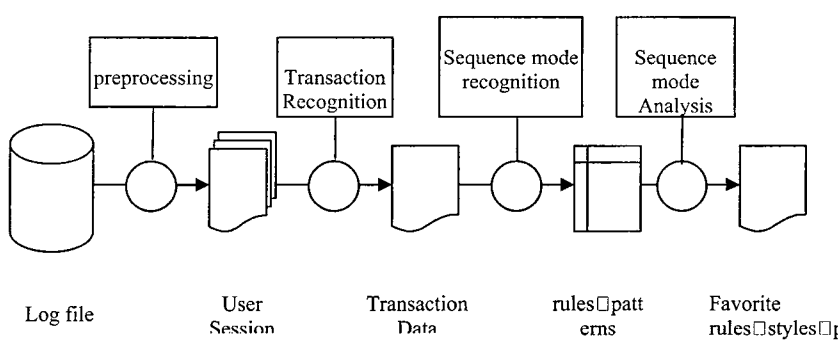

Figure 2: Logical processing consequence of log mining.

The steps of program processing are shown in Figure 3.

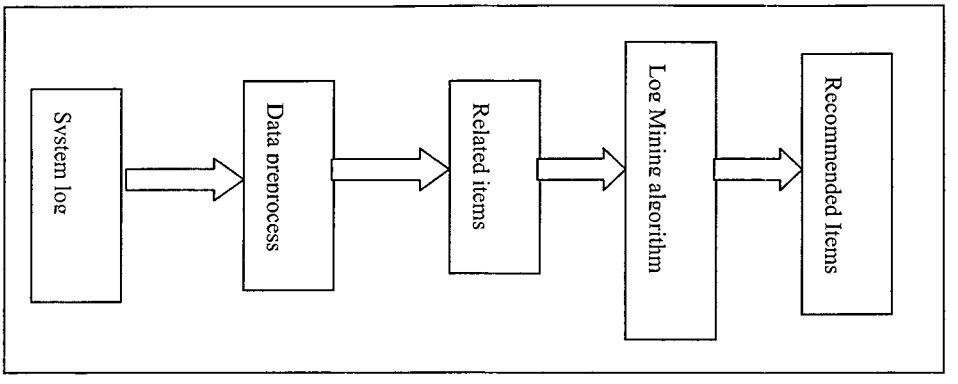

Figure 3 :Steps of program processing.

THE IMPLEMENTATION OF SYSTEM LOG MINING

First of all, we know the following three processed results of log have great impact on the new contract items: the latest used contract items from the log; the most frequently-used contract items; the most similar transaction contract items. The above-mentioned three sorts of contract items are most useful and likely. It is obvious that the contract item, which has been used in the shortest latest waiting time, is highest in preference degree. The more the contract items are used by aimed users, the higher the preference degree is and the more they are referred to (the higher support degree), the higher probability that the contract items is referred to. We analyse the visiting patterns and the preference structure of the aimed users, excluding the impact upon them by others. So we can ignore the third sort of contract items. Therefore, preference degree is the function of latest waiting time and frequency of referring, in which preference degree (pd) is negatively correlated with the latest waiting time (lwt) and positively correlated with the usage frequency (f), we can put it in this formula:

$$pd=\alpha\times(K-lwt)+\beta\times f \quad (1)$$

Here, α and β are the weighted coefficients, K is a constant.

Because lwt and f are the only two factors effecting pd, so weighted coefficients $\alpha+\beta=1$, and the formula can be simplified to

$$pd=\alpha\times(K-lwt)+\square 1-\alpha\square\times f \quad (2)$$

Here lwt is the number of days between the latest referred day of the aimed contract items and today; f is the referred times of the contract items during the past 3 years, so K is the total days of the past 3 years, so K should be $K=365\times3=1095$.

$$pd=\alpha\times(1095-lwt)+(1-\alpha)\times f \quad (3)$$

Obviously, $\alpha>0.5$, that means influence of frequency is less then that of lwt, that is the probability of the most frequently-used contract items during the past few years is no bigger than that of the latest items in the contract signed yesterday.

The coefficient α can be estimated through the employees' survey, we can begin with an initial weighted coefficient $\alpha=0.65$; The users may adjust it according to his/her practical situation. And the adjustment of contract items every time will enter into the log table as the basis of the recommended contract items of next time. The structure of log table is shown in table 1.

Now the preference degree formula becomes:

$$Pd=0.65\times(1095-lwt)+0.35\times f \quad (4)$$

THE APPLICATION OF SYSTEM LOG MINING

The difficulty does not lie in the realization of self-defined interface color, background images and sound but in the uncertain demand for the system content proposed by customers, which makes it impossible for us to predict their operating content for sure, in addition to their different operating habits, different professions and different operating contents.

Field name	Field Meaning	Field Type	Length	Comment
Logtype	Type of log	Varchar2	40	Example□Contact log,plan log,bid log,supplier log etc.
Logindex	Index number of log	Char	20	Example□20050321H21223
Logtime	Time of log	DateTime		Example□2005-03-21 14:42:09
LogIP	Source IP of user	Char	30	Example□192.168.6.127
LoguserID	User index number	Char	8	Example□41160004
Logusername	User's real name	Char	20	Example□Mr.Zhang
Refindex	Refer index	Char	20	Related index number of log Example□When logtype is contract□then refindex should be contract number, H41152005032100004
LogSQL	SQL statement	Varchar2	400	This field used for record the SQL statement when operation on database.
logclass	Class of log	Char	10	Example□normal□Danger□DataError□Warning
Comments	Description of log record	Varchar2	600	Example:this record created in the line 382 when contract printing.

Table 1: The structure of log table.

We take the realization of ERP's subsystem of contract management for a steel-making company as an example. There are fuel section, charge material section, scrape steel section, engineering section, mechanical spare parts section, electric appliance and instrument section, material section and so on in this company. Different sections are in charge of different materials.So the contracts have been classified into three types: Raw material and Energy contract, spare parts contract,material contract.Even in the same section, the employees have different work. For example, some employees in material section are in charge of chemical products; some building materials. In such circumstances, the quality standard item in contract is naturally different. While the user's personal demand is such that the most likely contract items will automatically be organized into a contract format, which can vary according to their different needs. According to the traditional method, we should design a program or templates for each of the customers. If the contents of employee jobs are redefined, the original system can no longer suit to the new conditions automatically. Therefore the system log must be used to mine users' personal needs and to match the content with their needs automatically.

The situation in contract printing is ssimilar: there is no fixed rule to print certain contract items. And the only way we can do it is to mine the user's preference information through the system log and to recommend the proper printing form.

CONCLUSIONS

Mining the users' preference information through the system log of ERP system can lead us to the construction of a intelligent personalized ERP system. The feature of this system log mining has won the praise of users. According to a statistics: over 90% of the recommended forms meet the user's needs. As to the others, a few adjustments to it will do. One thing that need to be improved is there is no comparison of similarity degree between different contract items. Currently, in recognizing the contract items, the method we use is to see if the characters are completely matching. It would be better to use the degree of similarity, which could be an important future research area.

REFERENCES

Anand S S , Patrick A R , Hughes J G (1998). A data mining methodology for crosssales. Knowledge Based Systems Journal , 10 (7) :449□461

Bfichner A G, Baumgarten M, Artand S S (1999). Navigation pattern discovery from internet data. In : Proceedings of the 5th ACM International Conference on Knowledge Discovery and Data Mining (WEBKDD'99 Workshop) (SIGKDD'99) , New York , 1999. 25□30

DongShan Xing□JunYi Shen QianBao Song (2003), Minging the users'preference directory from the web log (J), Transaction of Computer□26□11□:1518□1523

Park J S , Chen M S , Yu P S (1997). Using A hash2based method with transaction trimming for mining association rules. IEEE Transactions on Knowledge and Data Eng. 9(5) :813□825

Srikant R , Agrawal R (1995). Mining generalized association rules. In : Proceedings of the 21st International Conference Very Large DataBase ,Switzerland ,. 407□419

Srikant R , Agrawal R (1996). Mining quantitative association rules in large relational tables. In : Proceedings of the ACM SIGMOD , Canada , 1996.1□12

XinZhong Chen, Yan Li□YongHong Xie□BingRu Yang (2002). Web MiningResearch (J)□Engineering and Application of Computer, ,□13): 42□44

YingJi Li, Hong Peng, QiLun Zheng (2004). Web log clustering basis on the user's tasks (J), Mini-micro Computer System, ,(9):1620□1623