

Enhancing Smart Home User Behavior with Improved Data Mining Techniques

Miss. Pranjali Channe
Wainganga College of Engineering and Management
Dongergaon Nagpur, India

Abstract—Now in smart home system and abroad there found a wide variety of household appliances. These appliances are mostly used by the user and thus accumulating huge amount of life related data. This data contains wide information about user behaviour. Collecting this data and mining it for user behaviour analysis is now creating more interest in researchers. This paper approaches the improved techniques for data mining and focuses on reduction of delay of user behaviour analysis using these improved data mining techniques.

Keywords-Data Mining; Association Rule; User behaviour Analysis

I. INTRODUCTION

A smart home system has always under attention of global IT people and home appliances manufacturers. A fully fledged smart home system will contain a wide variety of household appliances. With the improvement in living standards of people, smart home is becoming the next standard of home life. Smart home not only gives users a safe, healthy and comfortable living environment, but also helps users to remotely monitor their home state with control home appliances. [6] Many people start to pay attention to make full use of the data processing capabilities of smart home devices to analyze data of the smart home appliances to extract the user's behavioural patterns of life and their habits, and finally provide users with personalized service and remind.

Association rule is the mostly used data mining technique. As an important part of the data mining technology, association rule mining aims at investigating how to find out the underlying rules and links through the massive amount of data. For this module input data should be gathered from all available sources, and the results of analysis will be analyzed for all possible goals. The collected data could be in scattered form, considering the diversification of those data that generated from our daily life, the focus of work is on how to use these existing data to analyze users' behavior and how to provide personalized services to different people under the specific conditions such as a particular time or place. [1]

This method could be helpful for the decision-making. In this paper, based on the different Association rules, a series of deep research will be conducted including mining towards the massive data associated with our life and find out the potential information that users preferred.

The second part of this paper introduces literature review; The third part describes problem statement on which our work is going to progress; the fourth part introduces the detailed architecture of the system; The fifth part gives idea about the possible contribution in; The last part summarize the text.

II. LITERATURE REVIEW

In data mining, association rule mining is the fundamental task and it is the most common method of data mining. Already there are many studies analyzing user behavior with association rule algorithm which are successfully applied in a

variety of scenarios. However the work done previously is based on the existing data and it will further refine the data mining association rules and adapt it to the isomerization data mining in the life data from smart home, finally find out the characteristics of user behavior. In this paper our work is subjected to enhance the existing data mining techniques.

A. User Behavior Analysis Models

An universal analysis model in mining is usually developed by following steps: data collection, data processing, data mining and analysis. Recently available user behavior analysis system has many applications on Internet technology, such as recommender system and the advertising of Google and Tencent. [1] The recommender system can find out the hobbies of user by analyzing the characteristics of website and analyzing the history traces of users' actions. And combines users' interesting with marketing strategies. A good model needs to meet these demands:

- 1) The subdivision of requirements;
- 2) The database processing capacity and fault tolerance capacity;
- 3) The strong adapt ability.

B. Apriori Algorithm

Apriori algorithm is a classical algorithm using association rules in data mining. Let $I = \{i_1, i_2, \dots, i_n\}$ be a set of data items with n different items. Given a transaction database D , there are m transactions and n data items in database D , the transaction T in D made by a subset of I .

Assume that $L \subseteq I$, if $L \subseteq T$, we claimed that the transaction T contains L . If the number of elements in the L is k , then L is k -item sets [1].

Scan database for first time, calculate all the individual items' support degree, compose one-dimensional frequent item sets (L_1) with items which support degree are not greater than minimum support degree. Then repeat the scan for k times, and connect K -frequent item (L_K) sets with each other to generate $K+1$ -frequent item sets, loop this process until no longer new frequent item sets appeared. The advantage of Apriori algorithm is pruning item set effectively by not generating and counting those item sets which could not become frequent item candidate set. Apriori employs an

iterative approach known as a level-wise search, where k-item sets are used to explore (k+1)-item sets. To improve the efficiency of the level-wise generation of frequent item set, the Apriori property (i.e., all nonempty subsets of a frequent item set must also be frequent.), an important property, is used to reduce the search space.

In general, association rule mining can be viewed as a two-step process:

- 1) Find all frequent item set;
- 2) Generate strong association rules from the frequent item set

III. PROBLEM STATEMENT

The problem definition is in the normal home, there are a variety of home appliances that can represent the behavior of the user.

The smart home is a home that predicts the nature of human activity that he/she performs on electronic devices with the help of information gathering in order to predict his/her behavior according to that information.

Here, the problem is in gathering that information from the number of devices into the single device and predicting human nature as it is difficult to collect all information and make the data list according to that on which human behavior will be predicted and also deciding the factors which will be useful for classifying human behavior.

IV. SYSTEM ARCHITECTURE

The basic user behavior analysis model consists of four parts: data acquisition, data preprocessing, data mining, and analysis of results.

Data acquisition will collect data from different data sources that are available in smart homes and data preprocessing will integrate and normalize data that come from different terminals which leads to improve the quality of data mining. In the part of data mining, we will analyze the data by means of association rules, and other improved data mining techniques to get predictable results ultimately in short time.

The detailed diagram of system architecture is shown in Figure 1:

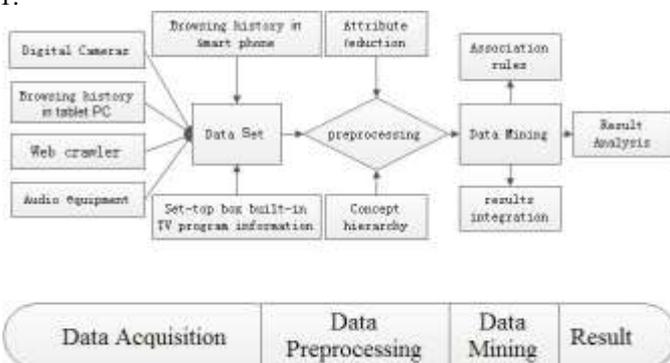


Figure 1. The detailed diagram of system architecture

V. POSSIBLE CONTRIBUTION

In this paper we proposed a model for user behavior analysis with the help of improved data mining techniques. However,

using traditional data analysis techniques to deal with the collection of data sets and their mining does have many limitations.

In this paper the two improved techniques are used to enhance the current system of behavior analysis for the smart home. To enhance the current system of user behavior analysis here we implement enhanced data mining techniques like Top K Rules and MOPNAR to improve the rule mining for smart home. The focus of using these improved techniques is to speed up the analysis process.

C. Top K Rule

Top k association rule has been developed to overcome the problem in association rule which leads to decide the user how to choose the thresholds to generate a desired amount of rules.[4]

Depending on the choice of the thresholds, current algorithms can become very slow and generate an extremely large amount of results or generate none or too few results, omitting valuable information.

To solve this problem, we propose to mine the top-k association rules, where k is the number of association rules to be found and is set by the user.

D. MOPNAR

This technique is used to remove the duplicates and inconsistency from the datasets. When the pre-processing on dataset is done the datasets are ready for evaluation. On the pre-processed datasets mining algorithms are applied.[2]

MOPNAR allows us to mine reduced sets of interesting, which provide us with interesting knowledge of the whole datasets, presenting average values of coverage greater than 99%.[2]

To improve resource utilization we combine MOPNAR with Top k Rules algorithm to devise a new algorithm BMPNAR i.e. Best M positive negative Association rules algorithm.[3] The proposed method is an extension to MOPNAR.

VI. CONCLUSION

This paper introduces improved data mining techniques to enhance the mining of huge data collected from different data sources of a smart home system. Thus association rule mining model for isomerization life data from multiple data sources tends to increase its analyzing performance.

REFERENCES

- [1] Bo Yin, Zhe Zhang, Xi Wang, Zhiqiang Wei "Research and Application of Data Mining Technology Used in the Analysis of Smart Home User Behavior" 2014 Sixth International Conference on Measuring Technology and Mechatronics Automation IEEE 2014
- [2] Ms. Sheetal Naredi, Mrs. Rushali A. Deshmukh, "Extraction and Classification of Best M Positive Negative Quantitative Association Rules", International Journal of Scientific & Engineering Research, Volume 6, Issue 7, ISSN 2229-5518, 2015
- [3] Ms Sonam Sheikh, Prof. Abhijeet Bajpayee, "Predicting User Behavior using Enhanced Data Mining Approach for Prediction in Smart Home", International Journal on Recent

- and Innovation Trends in Computing and Communication
Volume: 3 Issue, ISSN: 2321-8169 2 331– 334, 2015
- [4] Philippe Fournier-Viger, Cheng-Wei Wu and Vincent S. Tseng, "Mining Top-K Association Rules", University of Moncton, Canada, National Cheng Kung University, 2010
- [5] Amit Badlani, Surekha Bhanot, "Smart Home System Design based on Artificial Neural Networks," Proceedings of the World Congress on Engineering and Computer Science 2011 Vol I WCECS 2011, San Francisco, USA, 2011
- [6] Subhash K. Shinde, Dr. U. V. Kulkarni, "A New Approach For On Line Recommender System in Web Usage Mining", International Conference on Advanced Computer Theory and Engineering, 2008
- [7] Diana Martín, Alejandro Rosete, Jesús Alcalá-Fdez, "A New Multiobjective Evolutionary Algorithm for Mining a Reduced Set of Interesting Positive and Negative Quantitative Association Rules", IEEE transactions on evolutionary computation, vol. 18, no. 1, February 2014
- [8] Agrawal R, Imielinski T, Swami A. Mining Association Rules between Sets of Items in Large Databases [C]. In: Proc of the ACM SIGMOD International conference on Management of Data, Washington DC, 1993: 207-216
- [9] Agrawal R, Srikant R. Fast algorithms for mining association rules [A]. Proc. 1994 Int'l Conf. Very Large Data Bases (VLDB'94) [C]. Santiago: 1994: 487-499.
- [10] Ball, Robert, and Chris North. Analysis of user behavior on high-resolution tiled displays. Human-Computer Interaction-INTERACT 2005. Springer Berlin Heidelberg, 2005: 350-363
- [11] Borgelt C, Kruse R. Induction of association rules: Apriori implementation [C] // Compstat. Physica-Verlag HD, 2002: 395-400.