ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

Analyze Abnormal User Behavior Using Time sensitive Queries

[1]Parasa Gayatri, [2]J.Srinivas Rao
[1]Final M Tech Student, [2]Professor,
[1,2]Dept of Computer Science and Engineering, Nova college of Engineering & Technology, Jupudi, Ibrahimpatnm, Vijayawada.

Abstract: Now a day's underlying the searching process for user histories in real time oriented web applications is the main aspect in web mining. For understanding this problem, traditional approaches Evolving Agent Behavior classification based on Distribution of relevant events (EVABCD). It is based on representing the observed behavior of agent as an adaptive distributed of relevant atomic behaviors. Using this approach creating and updating user profiles from stored users in related aspects presented in database. Although EVABCD has been developed to be used online, the experiments have been performed using a batch data set in order to compare the performance to established (incremental and nonincremental) classifiers. The underlying assumption in this approach is that the data collected from the corresponding environment can be transformed into a sequence of events. But in traditional approach we are not evaluating the abnormal user behaviors with time varying queries. To address this problem we are introducing Top-K queries for retrieving user updates automatically. Our proposed work is used to develop evaluating the user profiles with invariant time sensitive attributes. It is an efficient process for retrieving relevant information of abnormal user profiles automatically.

Index Terms: *fuzzy-rule-based (FRB) classifiers, user modeling, time-varying, digital communications.*

I. INTRODUCTION

Now a day's recognizing the behavior of users in real time is a significant aspect of different human tasks in different environments. [1] [2] This process was carried out by using some software agents like robots based on User Modeling. [1] The recognition of users can be very beneficial for assisting them or predicting their future actions. Considering these results we are creating every user perspective. Computer user modeling is the process of learning users by observing the way they use the computer. This process needs the creation of a *user profile* that contains information that characterizes the usage behavior of a computer user. Computer users, like all of us, leave out activities that they do not even notice they are doing. Using of robotic systems to be successful in complex domains, but it is not useful in simple oriented applications they must be able to adapt to the environment, especially to the current behavior of other agents. [2] [3] User profiles can be description of the user interests, characteristics, behaviors, and preferences. User profiling is the practical orientation of gathering, organizing, and interpreting the user profile information. In recent years, significant work has been carried out for

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

profiling users, but most of the user profiles do not change according to the environment and new goals of the user.

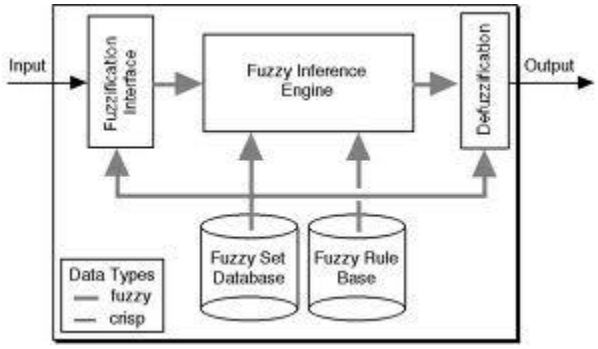

Figure 1: Fuzzy rule based searching process.

The above diagram shows the relevant information gather from our stored fuzzy set database. Our traditional approach does not capture the behaviors that cloud appears in the data stream once the classifier is built. The relevant information produced by users is often very large assessment for doing operations in the commercial accurance in the command execution in real time process execution. [4] [3] A complex task was introduced for retrieving relevant information about user behavior from database. Therefore, we need to cope with large amounts of data and process this information in real time, because storing the complete data set and analyzing it in an offline (batch) mode would be impractical.

II. RELATED WORK

Navigation on the World Wide Web relies on two main technical bases: selection of hypertext links or queries on search engines. [3] [6] While each technique constitutes a different paradigm, both are interesting in defining links to support user navigations over information. In recent years, significant work has been carried out for profiling computer users. In this research, an approach for profiling and recognizing *general* user behavior profiles is proposed. Different techniques have been used to find out relevant information related to the human behavior in many different areas. The literature in this field is vast; Macedo et al. [1] propose a system (WebMemex) that provides recommended information based on the captured history of navigation from a list of known users. Pepyne et al. [1] describe a method using queuing theory and logistic regression modeling methods for profiling computer users based on simple temporal aspects of their behavior. In this case, the goal is to create profiles for very specialized groups of users, who

would be expected to use their computers in a very similar way. [7] Popular approaches to such learning include statistical analysis and frequency based methods.

Several works have demonstrated the appropriateness of the open hypermedia approach, and a protocol has been proposed towards allowing the interchange of information among applications. In this paper, we propose an application which is continuously recording users' Web surfacing activity and using it to automatically define links between related documents visited. [3] Thus, it is necessary that the approach deals with the problem of classification of streaming data. Incremental algorithms build and refine the model at different points in time, in contrast to the traditional algorithms which perform the model in a batch manner. It is more efficient to revise an existing hypothesis than it is to generate hypothesis each time a new instance is observed. Therefore, one of the solutions to the proposed scenario is the incremental classifiers.

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

Using these approaches we are developing efficient algorithms for constructing data retrieval techniques.

III. EXISTING SYSTEM

The novel evolving user behavior classifier is based on fuzzy system and it takes into account the fact that the behavior of any user is not fixed but rarely is it changed. Although existing approach can be developed in the behavior represented by the details of user behavior profile represented in the sequence of events using some credentials present in the databases. When this process is carried out by software agents or robots, it is known as user modeling. The recognition of users can be very beneficial for assisting them or predicting their future actions. Most existing techniques for user recognition assume the availability of handcrafted user profiles, which encode the a-priori known behavioral repertoire of the observed user. However, the construction of effective user profiles is a difficult problem for different reasons: human behavior is often erratic, and sometimes humans behave differently because of a change in their goals. This last problem makes necessary that the user profiles we create evolve.

The traditional approach includes at each step the following two main actions:

Creating and evolving the classifier. This action involves in itself two sub actions:

a. Creating the user behavior profiles. This sub action analyzes the sequences of commands typed by different windows users online (data stream), and creates the corresponding profiles.

b. Evolving the classifier. This sub action includes online learning and update of the classifier, including the potential of each behavior to be a prototype, stored in the EPLib.

User classification: The user profiles created in the previous action are associated with one of the prototypes from the EPLib, and they are classified into one of the classes formed by the prototypes.

IV. PROPOSED APPROACH

In our approach we are introducing new technique EVABCD can also be used to monitor, analyze, and detect abnormalities based on a time-varying behavior of same users and to detect masqueraders. [2] It can also be applied to other type of users such as users of e-services, digital communications, etc. EVABCD needs an appropriate subsequence length to get a classification rate similar to the obtained by other classifiers which use different techniques. However, EVABCD does not need to store the entire data stream in the memory and disregards any sample after being used. [5] [2] EVABCD is one pass (each sample is preceded once at the time of its arrival), while other offline algorithms require a batch set of training data in the memory and make many iterations. Thus, EVABCD is computationally more simple and efficient as it is recursive and one pass. Unlike other incremental classifiers, EVABCD does not assume a prefixed structure and it changes according to the samples obtained. [4] In addition, as EVABCD uses a recursive expression for calculating the potential of a sample, it is also computationally very efficient. In fact, since the number of attributes is very large in the proposed environment and it changes frequently, EVABCD is the most suitable alternative.

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

EVABCD have following structure for evaluating the user profiles.

1. **Classify the new Sample:** It was describe the sample prototype for different users for maintaining their histories.
2. **Calculate Potentials:** Every sample data set can be maintaining newly search user details with new prototypes. [2]
3. **Update:** If any modifications are present in the newly coming user profiles. [3] We are also maintaining new prototype for storing that information with automatic consistency.
4. **Remove:** Remove the unnecessary results from old prototype for storing newly coming datasets.
5. **Supervised and Unsupervised Learning:** In this requirement we are assigning the prepare dataset for storing relevant information from relational dataset. In data sets representation we are formed the training data for preparing new user profiles based on their behaviors.

Excluding this environment we also add the time sensitive queries for retrieving relevant information of users with invariant time. [7] Time is an important dimension of relevance for a large number of searches, such as over blogs and news archives. So far, research on searching over such collections has largely focused on locating topically similar documents for a query. Unfortunately, topic similarity alone is not always sufficient for document ranking.

V. PERFORMANCE ANALYSIS

In this section we describe efficient results for our time varying queries present in the existing approach EVABCD. [3] We are evaluating the concurrent results for every user present in the data base.

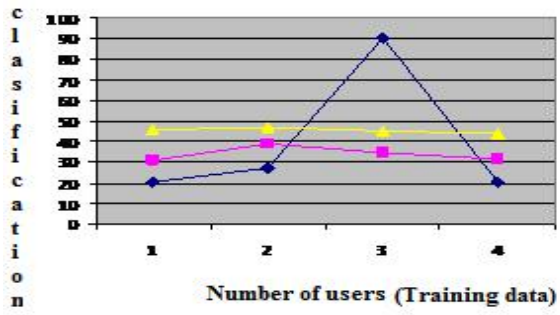

Figure 2: Clasification rate using data set.

Above diagram shows the four graphic results of this experiment considering in each graph one of the four classes as the

New class:

. X-axis represents the number of [5] users of the new class that contains the training data set.

. Y-axis represents the percentage of users of the new Class correctly classified.

In the different graphs, we can see how quickly EVABCD evolves and adapts to the new class. [2] If we consider the class Novice Programmers, it is remarkable that after analyzing three users of this class, the proposed classifier is able to create a new prototype in which almost 90 percent of the test users are correctly classified. However, the other classifiers need a larger number of samples for recognizing the users of this new class. Similar performance has been observed for the other three classes. Specially, C5.0 needs several samples for creating a suitable decision tree. As we can see in the graph which represents the class Computer scientist as the new class, the percentages of users correctly in the 1-NN classifier is always 0 because all the users of this class are classified in the Novice programmers class. The increase in the

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

classification rate is not perfectly smooth because the new data bring useful information but also noise. [2] Taking into account these results, we would like to remark that the proposed approach is able to adapt to a new user behavior extremely quick.

We show that our techniques are robust and significantly improve result quality for time-sensitive queries [7] compared to state-of-the-art retrieval techniques.

VI. CONCLUSION

Creating the user behavior profiles. This sub action analyzes the sequences of commands typed by different windows users online (data stream), and creates the corresponding profiles. [4] [1] The test results with a data set of 168 real UNIX users demonstrates that, using an appropriate subsequence length, EVABCD can perform almost as well as other well-established offline classifiers in terms of correct classification on validation data. Our proposed work is used to develop evaluating instead queries of the user profiles in real time oriented top entry abnormal user histories. It is an efficient process for retrieving relevant information of abnormal user profiles automatically. [7] Our work demonstrates that integrating time in the retrieval task can improve the quality of the retrieval results, and motivates further research in the area. Currently, we rely on the publication time of the documents to locate time periods of interest. [7] Further enhancement for our proposed work is to introduce time-based diversity in query results by grouping the results into clusters of relevant time ranges, enabling users to be aware of and interact with time information when examining the query results.

VII. REFERENCES

1) D. Godoy and A. Amandi, “User Profiling in Personal Information Agents: A Survey,” Knowledge Eng. Rev., vol. 20, no. 4, pp. 329 361, 2005.

2) Jose Antonio Iglesias, Member, IEEE Computer Society, Plamen Angelov, Senior Member, IEEE Computer Society, Agapito Ledezma, Member, IEEE Computer Society, and Araceli Sanchis, Member, IEEE Computer Society” Creating Evolving User Behavior Profiles Automatically”, IEEE, 2012.

3) Jos´e Antonio Iglesias, Agapito Ledezma, and Araceli Sanchis” Creating User Profiles from a Command-Line Interface: A Statistical Approach”, IEEE, 2011.

4) Alessandra Alaniz Macedo, Khai N. Truong, Jos´e Antonio CamachoGuerrero, Maria da Grac¸a Pimentel” Automatically Sharing Web Experiences through a Hyper document Recommender System”,IEEE,2010.

5) Patrick Riley and Manuela Veloso” On Behavior Classification in Adversarial Environments”, IEEE, 2010.

6) José Antonio Iglesias, Agapito Ledezma, and Araceli Sanchis” A Comparing Method of Two Team Behaviours in the Simulation Coach Competition”, International Conference, 2010.

7) Wisam Dakka, Luis Gravano, and Panagiotis G. Ipeirotis “Answering General Time-Sensitive Queries”, IEEE Transactions On Knowledge And Data Engineering, Vol. 24, No. 2, February 2012.