International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-12)

Research Article | **December 2015**

A Comparitive Study of KNN and EvABCD Classification Method for Customer Behavior

[1]Dr. P. Sumathi, [2]B. Uma Maheswari
[1]Asst. Professor, Govt. Arts College, Coimbatore, Tamil Nadu, India
[2] Research Scholar in Bharathiyar University, Coimbatore, Tamil Nadu, India

Abstract:

E-commerce is the new business techniques which deal with the various needs of organizations, retailer and consumers to reduce cost by developing the goods quality and services which also increase the delivery speed with the help of Internet. This is different from the established electronic commerce by enabling the goods trading, money and electronically information from computer to computer. Information about the users in computers is helpful in forecasting the future behaviors of the computer user. An Evolving Agent Behavior Classification based on Distributions (EvABCD) method mainly depends on the customer's observed behavior. This will modify the user's interest in search specified in the user profiles automatically when the user interest in search changes and will retrieve the search results on internet. This new method is very efficient in adapting the evolving user profiles and provides the better search results based on the user's interest.

Keywords: E-commerce, k-NN, Evolving Agent Behavior Classification based Distribution, Interested users

I. INTRODUCTION

A most important and popular web-based activity is an E-Commerce; it may perform different types of financial transactions or facilitated using the Web. It is generally estimated that E-Commerce activity will perform function to develop and that it will be an important component of the global economy in the near future. A significant issue in designing E-Commerce systems is to characterize the customer's requirements for reasonable service. Parameters which change a customer's satisfaction with an E-Commerce system contains the response time, amount of clicks needed to find what they want, huge amount of information are required to give, and predictability of the service received. This leads to the idea of customer classification, where customers in the similar class would value parameters in a same fashion. Customer classification may do the operation of either based on how they judge their satisfaction with an E-Commerce system, or on some another way (e.g. type/speed of Internet connection the customer has to the server; frequent/previous/new customer; large/medium/small budget).

A fundamental component in search engine personalization is a good user profiling strategy. Several personalization methods is related on the construction of one single profile for a user and user have to indicate his search interest in that and based on that interest, the results will be retrieved. If search interest changes, have to update that manually. Different queries from the user should be handled differently because a user's preferences may change across queries. Based on the changing behavior of the user, profiles should be updated automatically.

The problem is that it requires too much insight understanding and effort from the user to recognize the opportunity to employ an agent take the initiative to create it supply the agent with explicit knowledge and maintain the underlying rules or scripts over time. The basic assumption in this technique is that the data's are grouped from the corresponding environment can be converted into a sequence of events. But in traditional approaches are not evaluating the abnormal user behaviors with time varying queries. To address this problem introducing a new approach called Evolving Agent Behavior Classification based on Distributions of relevant events. An adaptive technique for generating behavior profiles and recognizing various computer users called Evolving Agent Behavior Classification based on Distributions of relevant events is used and it is based on representing the observed behavior of an customer as an adaptive distribution of her/his relevant atomic behaviors (events).

This paper is summarized as follows: an overview of related work is given in Section 2. The main performance of the proposed technique is provided in Section 3. Experimental results are discussed in section 4. Conclusion and future work provided in section 5.

II. LITERATURE SURVEY

Profiling the behavior of programs can be a useful reference for detecting potential intrusions against systems. Ghosh *et al* (1999) presented three anomaly detection techniques for profiling program behavior that evolve from memorization to generalization. Thus, it is essential to take into account these changes in any behavior recognition system. A general approach to the classification of streaming data which represents a specific agent behavior based on evolving systems are given by Iglesias *et al* (2010). An approach for discovering and tracking evolving user profiles are proposed. Nasraoui *et al* (2008) described how the discovered user profiles can be enriched with explicit information need that is inferred from search queries extracted from Web log data.

A user profile is not necessarily fixed but rather it evolves/changes, Iglesias *et al* (2012) proposed an evolving method to keep up to date the created profiles using an Evolving Systems approach. They combined the evolving classifier with a trie-based user profiling to obtain a powerful self-learning online scheme. Benevenuto *et al* (2009) presented a first of a kind analysis of user workloads in online social networks. Viswanath *et al* (2014) proposed using unsupervised anomaly detection techniques over user behavior to distinguish potentially bad behavior from normal behavior. Jin *et al* (2013) suggested a comprehensive review of state-of-the-art research related to user behavior in OSNs from several perspectives. First, they discussed social connectivity and interaction among users. An approach used for creating and updating automatically the profile of a computer user is used called evolving agent behavior Classification based on distributions of relevant events by Sushma and Ramesh (2013).

Dharani and Geetha (2013) suggested a user profile which is evolved using Case Based Reasoning. The learner's behavior of the E-learning system is tracked dynamically using the Colored Petri Nets. The data mining techniques and an application to contributed in e-commerce for completes the web personalization is proposed by Gawali *et al* (2013). Personality traits are mainly measured by psychological questionnaires, and it is hard to obtain personality traits of large amount of users in real-world scenes. Gao *et al* (2013) proposed a new approach to automatically identify personality traits with Social Media contents in Chinese language environments. Lacerda and Nivio (2013) analyzed different aspects of user profiling: selecting the most informative events from the interaction between users and the system, and combining different recommendation algorithms.

There have been few efforts on understanding the correlations between users' social media profiles and their e-commerce behaviors. A system for predicting a user's purchase behaviors on e-commerce websites from the user's social media profile is presented by Zhang and Marco (2013). Customer loyalty or repeat purchasing is critical for the survival and success of any store. By focusing on online stores, Chiu et al (2014) investigates the repeat purchase intention of experienced online buyers based on means-end chain theory and prospect theory. Rodriguez and Angel (2011) presented an attempt to analyze the factors that determine e-commerce adoption by final consumers. The proposed theoretical model is applied on two different samples: one composed of internet users with no previous experience of virtual shopping, and another formed by subjects that have already made online transactions previously.

The rapid spread of e-commerce has created tremendous opportunities for economic efficiency and customer choice. Even many consumers and businesses are reveling in e-commerce; consumer problems related to online retail become the dark side of the issue. This study motivates the e commerce by rectifying the problem using EvABCD algorithm.

III. THE PROPOSED APPROACH

Electronic commerce refers to the buying and selling of information, products and services through computer network. This section introduces the proposed approach for classification of the behavior profiles of customers. Due to changing in the customer's behaviour, it is not rigid. But this proposed approach, can be applied to all behavior which is represented by event sequence, this chapter may use an interesting visitors record interface environment. The classifier of KNN and EVABCD is used to evaluate the performance of the customer.

3.1 K Nearest Neighbor

The K-nearest neighbour algorithm is a machine learning technique that belongs to the class of instance based learners. Given a training set, a similarity measure over patterns, and a number K, the algorithm predicts the output of a new instance pattern by combining the known outputs of its K most similar patterns in the training set. The training data is stored in memory and only used at run time for predicting the output of new instances. Advantages of this technique are the (implicit) construction of a local model for each new instance pattern, and its robustness to the presence of noisy training patterns. This k-NN algorithm is successfully applied on many pattern recognition problems i.e., scene analysis. The k-NN algorithm application is used to predict problems in call centers which translates to the prediction of the arrival rate function till the end of planning horizon, based on matching K arrival rate functions to the pattern observed so far.

More formally, suppose that it has observed the call arrival rate $r_1, \dots, r_x$ on a specific day in periods $t_1, \dots, t_x$ where $x < n$. Let us call this information the reference trace $\vec{r}_x = (r_1, \dots, r_x)$. The nearest neighbour algorithm compares the reference trace to historical data $\vec{h}_{d,x} = (h_{d,1}, \dots, h_{d,x})$ from the matrix H for $d = 1, \dots, m$. The comparison is based on a distance function $D(\vec{r}_x, \vec{h}_{d,x})$, which is defined by the norm on the space of the traces. The K nearest traces with respect to the distance function D is used to generate for the arrival rates $\hat{r}_{x+1}, \dots, \hat{r}_n$ in periods $t_{x+1}, \dots, t_n$. The result depends on the value of K as well as on the choice of the distance measure D.

Along with the development of computer network, the electronic commerce has become the new pattern to carry on the commercial activity gradually, but the security problem is also getting more and more prominent in KNN. To overcome from this problem proposed method of EVABCD is used.

3.2 Evolving Agent behavior Classification based on Distributions of relevant events (EVABCD)

The EVABCD approach for automatic classifier design of the behaviour profiles of users. Original evolving user behaviour classifier is based on Evolving Fuzzy Systems and it takes into account the fact that the behaviour of any user is not fixed, but is rather changing. In this paper, propose an adaptive approach for creating behavior profiles of the customer. It is based on representing the observed behavior of interesting visitors as an adaptive distribution of her/his relevant atomic behaviors (events). Once the model has been created, EVABCD presents an evolving method for updating and evolving the customer's interest. Thus the goal of EVABCD can divide into two phases:

1. Creating and updating user profiles from the records.
2. Classifying a new sequence of records into the predefined profiles.

This action involves in itself two sub actions:

- Creating the user behaviour profiles. This sub action will study the interesting visitor's records sequences typed by different users online and generate the related profiles.
- Classifier Evolution. The sub action consists of online purchasing and classifier update, also includes the possible behaviour to be a model and stored in EPLIB.
- Classification of user. The created web page visitor's profiles in the preceding section are related with the one of the prototype from the EPLIN and classify into classes created by the model.

EVABCD have following structure for evaluating the interesting visitor's 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 pages details with new prototypes.
3. Update: If any modifications are present in the web page. 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, assigning the prepare dataset for storing relevant information from relational dataset. In data sets representation is formed the training data for preparing new customer profiles based on their behaviors.

Thus, EVABCD is computationally more simple and efficient as it is recursive and one pass. In fact, since the number of attributes is very large in the proposed environment and it changes frequently, EVABCD is the most suitable alternative. Finally, the EVABCD structure is simple and interpretable to identify the interested visitors.

IV. EXPERIMENTAL RESULTS

In this section describe the efficient results for our time varying queries present in the existing approach EVABCD. It evaluates the concurrent results for every customers present in the data base. Experimental result is implemented by using MATLAB.

Table 1: Attributes used for e commerce

Attributes	Description
A1	Look for product offers
A2	Price details
A3	Read sub pages
A4	Product benefits
A5	Visit all pages for few minutes
A6	visit web page regularly

The characteristics of the attributes are discussed in table 1. It contains six attributes to find the interested and non interested buyers.

Table 2: Rate for interested and non-interested customer

Customer Class	Attributes	Rate
With purchase interest	A1	0,0,1,0,1,1
	A2	1,1,1,1,1,0
	A3	0,2,1,3,0,1
	A4	1,0,1,0,1,1
Without purchase interest	A5	0,1,0,0,1,0
	A6	0,0,0,0,1,0

Table 2 shows the values for with purchasing and without purchasing interest. The interested buyers may have maximum value of above one for all the attributes, but non interest buyers have maximum value as zero. Table 2 taken a six customer from huge amount of customers and show the result.

Table 3: Classification Rate (Percent) of Different Classifiers in the customer reading interest using Different Subsequence Lengths

Number of records for training	Classifier Rate (%)				
	Subsequences length	With purchase interest		Without purchase interest	
		K-NN	EvABCD	K-NN	EvABCD
100	2	32.6	20.2	19.3	20.1
	3	35.8	33.5	21.7	28.6
	4	36.3	65.8	33.5	40.9
	5	32.8	68.1	41.1	52.8

	6	37.7	70.4	45.6	65.7
500	2	38.2	73.7	43.8	65.9
	3	35.7	71.3	45.2	66.2
	4	39.8	73.8	45.8	67.4
	5	41.3	75.6	46.5	70.5
	6	43.8	75.9	49.3	71.3
1000	2	44.6	69.2	50.8	73.8
	3	44.9	75.0	51.2	74.6
	4	45.6	79.5	55.3	77.5
	5	46.2	82.7	56.8	79.3
	6	46.8	85.6	57.4	81.6

According to these data, EVABCD perform slightly better than the KNN classifiers in terms of accuracy. The percentages of users correctly classified by EVABCD are higher to the results obtained and lower than the percentages obtained by KNN. In general, the difference between EVABCD and the KNN is considerable for small subsequence lengths, but this difference decreases when this length is longer. These results show that using an appropriate subsequence length, the proposed classifier can compete well with offline approaches. Nevertheless, the proposed environment needs a classifier able to process streaming data in online and in real time. In addition, the learning in EVABCD is performed in single pass and a significantly smaller memory is used. Spending too much time for training is clearly not adequate for this purpose. Then the number of attributes is very large in the proposed environment and it changes frequently, EVABCD is the most suitable alternative. Finally, the EVABCD structure is simple and interpretable.

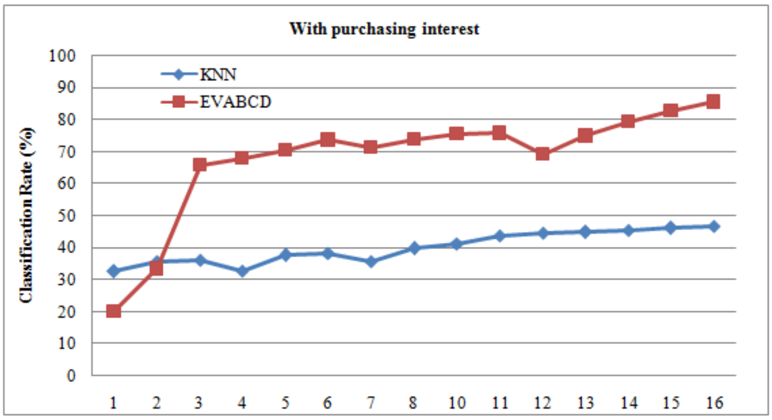

Figure 1: Classification rate for interested customer

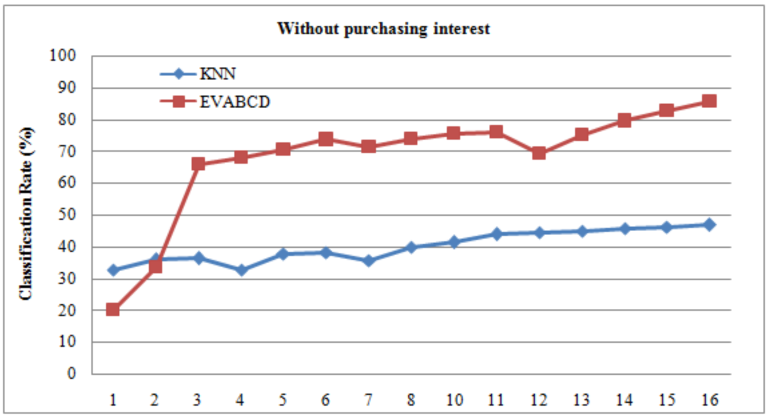

Figure 2: Classification rate for non-interested customer

Figure 1 and figure 2 show the interested and non-interested customers for purchasing in e-commerce. Thus the proposed method of EVABCD has better performance in classification rate when compare with the FNN.

V. CONCLUSION

Designing customer satisfaction with Web and E-commerce sites is not good study as web server modeling, but finding whether and how the customers of same sites are satisfied with their interaction is a most important same as web matures. A generic approach, EVABCD is implemented in E commerce system is used to classify and design user behaviors from an order of events. EVABCD is recursive, it is utilized in an interactive mode; consequently, it is computationally essential and high speed in updating the customer details. Therefore, its structure is easy and interpretable. This personalization technique can also be utilized to examine and detect abnormalities based on a time-varying behavior of similar users and it is used to detect interested/non interested users. Additionally, proposed approach is a fast retrieval process of search result; the classifier can be further learned and improved in future.

REFERENCES

[1] Benevenuto, Fabrício, Tiago Rodrigues, Meeyoung Cha, and Virgílio Almeida. "Characterizing user behavior in online social networks." In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pp. 49-62. ACM, 2009.

[2] Chiu, Chao‐Min, Eric TG Wang, Yu‐Hui Fang, and Hsin‐Yi Huang. "Understanding customers' repeat purchase intentions in B2C e‐commerce: the roles of utilitarian value, hedonic value and perceived risk." *Information Systems Journal* 24, no. 1 (2014): 85-114.

[3] Dharani, B., and T. V. Geetha. "Adaptive learning path generation using colored petri nets based on behavioral aspects." In *Recent Trends in Information Technology (ICRTIT), 2013 International Conference on*, pp. 459-465. IEEE, 2013.

[4] Gao, Rui, Bibo Hao, Shuotian Bai, Lin Li, Ang Li, and Tingshao Zhu. "Improving user profile with personality traits predicted from social media content." In*Proceedings of the 7th ACM conference on Recommender systems*, pp. 355-358. ACM, 2013.

[5] Gawali, Rajesh R., and Shivaji D. Mundhe. "Web Mining techniques, process and applications in Ecommerce." In *Proceedings of National Conference on Emerging Trends: Innovations and Challenges in IT*, vol. 19, p. 20. 2013.

[6] Ghosh, Anup K., Aaron Schwartzbard, and Michael Schatz. "Learning Program Behavior Profiles for Intrusion Detection." In *Workshop on Intrusion Detection and Network Monitoring*, vol. 51462. 1999.

[7] Iglesias, Jose Antonio, Plamen Angelov, Agapito Ledezma, and Araceli Sanchis. "Creating evolving user behavior profiles automatically." *Knowledge and Data Engineering, IEEE Transactions on* 24, no. 5 (2012): 854-867.

[8] Iglesias, Jose Antonio, Plamen Angelov, Agapito Ledezma, and Araceli Sanchis. "Evolving classification of agents' behaviors: a general approach."*Evolving Systems* 1, no. 3 (2010): 161-171.

[9] Jin, Long, Yang Chen, Tianyi Wang, Pan Hui, and Athanasios V. Vasilakos. "Understanding user behavior in online social networks: A survey." *IEEE Communications Magazine* 51, no. 9 (2013): 144-150.

[10] Lacerda, Anisio, and Nivio Ziviani. "Building user profiles to improve user experience in recommender systems." In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 759-764. ACM, 2013.

[11] Nasraoui, Olfa, Maha Soliman, Esin Saka, Antonio Badia, and Richard Germain. "A web usage mining framework for mining evolving user profiles in dynamic web sites." *Knowledge and Data Engineering, IEEE Transactions on*20, no. 2 (2008): 202-215.

[12] Rodríguez Del Bosque, Ignacio, and Ángel Herrero Crespo. "How do internet surfers become online buyers? An integrative model of e-commerce acceptance." *Behaviour & Information Technology* 30, no. 2 (2011): 161-180.

[13] Sushma, Y., and J. Ramesh. "Automatic Creation and Updation of User Behavior Profile." In *International Journal of Engineering Research and Technology*, vol. 2, no. 9 (September-2013). ESRSA Publications, 2013.

[14] Viswanath, Bimal, M. Ahmad Bashir, Mark Crovella, Saikat Guha, M. S. R. India, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. "Towards detecting anomalous user behavior in online social networks." In*Proceedings of the 23rd USENIX Security Symposium (USENIX Security)}*. 2014.

[15] Zhang, Yongzheng, and Marco Pennacchiotti. "Predicting purchase behaviors from social media." In *Proceedings of the 22nd international conference on World Wide Web*, pp. 1521-1532. International World Wide Web Conferences Steering Committee, 2013.