Journal of Internet Banking and Commerce

An open access Internet journal (http://www.arraydev.com/commerce/jibc/)

Journal of Internet Banking and Commerce, April 2013, vol. 18, no.1
(*http://www.arraydev.com/commerce/jibc/*)

Decision Support System for Electronic Multi-Banking (DSSEMB)

G SREE REKHA
Assistant Professor, MCA Department, PES Institute of Technology, Bangalore, India
Postal Address: **100 Feet Ring Road, BSK 3rd Stage, Bangalore - 560085, India**
Author's Personal/Organizational Website: www.pes.edu
Email: **sreerekha@pes.edu**

Ms. Sree Rekha is an Assistant Professor in Department of MCA, PESIT, Bangalore, India. Her areas of interest are Data Mining and pattern recognition, Web Banking Security, Information security, Architectural and protocol design & analysis.

DR. V. K. AGRAWAL
Professor ISE Department and Director R&D(CORI), PES Institute of Technology, Bangalore, India
Postal Address: **100 Feet ring Road, BSK 3rd Stage, Bangalore, 560085,India** *Author's Personal/Organizational Website:* www.pes.edu
Email: **vk.agrawal@pes.edu**

Dr. Agrawal is the Director CORI Lab (R&D) in the PES Institute of Technology, Bangalore, India. He is also a Professor in Information Science Engineering. His current research interests are Data Mining, Information security, Software and hardware architectures, Software engineering, embedded systems and Power systems. He has published many papers in various prestigious journals such as IEEE, International journal of computers and applications, International journal of embedded systems etc.

Abstract

Recent times have seen an explosive growth in the availability of various kinds of electronic banking for different purposes like e-shopping, e-trading, e-auction etc. In order to meet the challenges of the latest developments in the area, banks have to invest in new technologies. This will help in understanding the customers better and provide better services. This also helps in identifying the most powerful attacks, expected or unexpected frauds. Solutions for them could also be made possible.

Multi-banking is an environment in which a user can access multiple accounts with a single browser. To meet the customer's requirements like security, speed, reliability etc., new technologies have to be introduced by the banks. In this paper, we describe a prototype decision support system for electronic multi-banking. The Authentication server uses a Banking transaction Service Model for authenticating a user and notifying the banks either to provide or deny the services to a particular user. Banking Transaction Service model (BTSM) acts as a decision support system. The intermediate server acts as a link between the customer and various banks. We also briefly discuss an application of data mining for analyzing the large datasets collected from the information provided by the user.

Keywords: **Data Mining models, multi-banking, BTSM, Decision support system**

INTRODUCTION

Multi-banking is an environment which allows a particular user/customer to operate his/her multiple bank accounts, once he/she logged in to any bank's server[1]. The server of the bank has to act as a link between user and the banks until authentication is done. As the risk levels are high in multi-banking, it is always advisable to have an intermediate server for the purpose of authenticating a user. Again if the intermediate server is a third party, then the process could be much more complex. To the authentication server, all the credentials information is to be sent either for authentication or for service provision. We propose to use an Intermediate Authentication server (IAS) which is also a bank server, to authenticate a user and notify the banks regarding the request made by the user. The computations and calculations will be done by the IAS for decision making. For Example, a user logs on into the Bank-A and requested the services of Bank-B, Bank-C and so on, then the Bank-A acts as an Intermediate server between the user and the banks.

As an intermediate server the major activities performed by the Bank-A server are:

(a) Give the data provided and the data which is existed in the database to the model,
(b) Decision making and token generation for the genuine user,
(c) Notifying the banks to provide or deny the services. We propose to use a data mining model for the purpose of arriving at a decision.

Efficient utilization of data whichever is available could be helpful in the decision making process.

As the volume of data collected and stored in databases while performing online transactions are considerably more, one can think of using data mining techniques. Data mining is one such technique which is relatively a new trend[2]. Various companies and organizations adopt such methods in order to increase their profit and diminish their cost [3]. The development and continuous training of prediction models is a very significant task. Accurate prediction of future facts is possible if such models are applied to specific banking units. Data mining techniques enable successful prediction of various parameters related to banking. Prediction making requires the correct parameter determination that will be used for predictions and finally for decision-making.

Decision-making process in multi-banking is critical because of the sensitivity and complexity of data. Use of data mining models can be a solution for handling huge amounts of data to arrive at decisions. Data mining is a decision-support process, which is based on artificial intelligence (AI), machine learning, statistical, and other technologies. Highly automated analysis of the original enterprise data, forecasting the client's behavior and helping the decision makers is done by data mining [1, 2 and 3]. Adequate methods are needed to handle the most risky situations. The proposed system is designed to assist a bank to come out with a decision regarding provision/denial of services to a customer with respect to the request made by him/her who wishes to operate his/her multiple bank accounts with a single sign-in and it may represent as extension to the existing online banking activities.

Existing e-banking model: E-banking is an umbrella term for the process by which a customer may perform banking transactions electronically without visiting a brick-and-mortar institution [4]. As shown in the figure 1, in the existing e-banking model, the user has to start operating his/her bank account by entering into the particular bank's website through the internet. Then, the Internet banking server will receive the request through a firewall and process the request by verifying the data obtained with that already existing in the database.

Some of the major drawbacks in the existing system are:

A. A user can operate only one account at a time.
B. Information transmission is subject to many attacks.
C. Threat of Phishing attack is considerable.
D. No three factor (i.e., username, password and biometric) authentication.

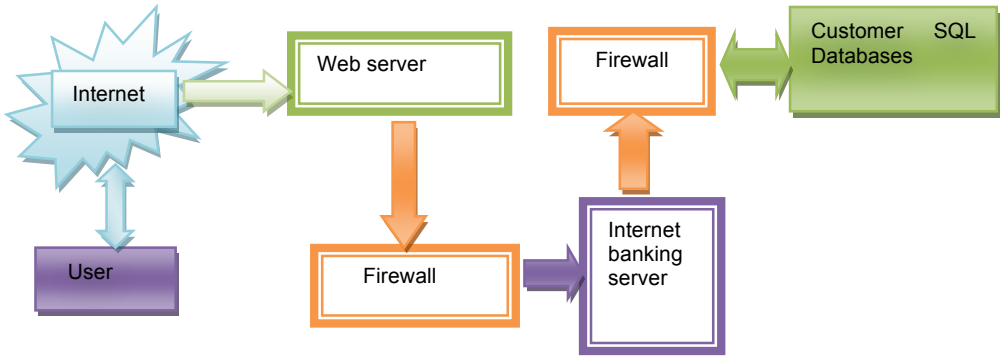


Figure 1: Process of Existing online banking system

The rest of the paper is organized as follows: Section 2 gives an overview of the proposed e-multibanking model and flow of messages. Section 3 describes the banking Transaction Service model (BTSM) for authentication and fraud detection. Section 4 gives an overview of generating the predictions for e-multibanking data set. Section 5 discusses the probability of success in hacking the existing system and in improvement using the data mining concept in the proposed system compared with the existing system. The last section deals with simulations and conclusion.

PROPOSED E-MULTIBANKING MODEL AND FLOW OF MESSAGES

In the proposed Electronic multi-banking model, a user initially sends a request to a bank. After successful completion of the authentication process, the bank server provides various options to enable multi-banking to the user. Then the bank server acts as an intermediary between the user and the other banks requested by the user. It is the initial bank server that gets the request from the user which acts as an intermediary server. And it is the sole responsibility of the server to authenticate a user. The Authentication server would use the data mining model for making the decision to provide or deny the services. Intelligent phishing detection system for e-banking using fuzzy data mining concept was earlier used in order to detect the phishing frauds [5].The following figure 2 gives a bird's eye view of the proposed model.

In the proposed e-multibanking model illustrated in figure 2, the participating entities are:

(a) A customer in the client side (C).
(b) An intermediate authentication server (IAS).
(c) Number of banks(B1, B2, B3 and so on, i.e. Bi for i-1,2,----------n where n is the number of banks) and
(d) Banking Transaction Service model (BTSM) which actually processes the data received intellectually by considering various parameters and comes out with a decision, which in turn would be communicated to the Intermediate Authentication Server (IAS).

The process which is to be carried out in the model presented in figure 2 is as follows:

I. Initially, the user sends a request to the bank requesting the services of other banks along with that particular bank's services.
II. The bank server to which the user has sent the request itself acts as an intermediate server for authentication and notification.
III. The IAS (Initial bank server to which the request is sent) comes out with a decision based on the model which considers various parameters dynamically, computes the result and supports the decision making process.
IV. Once the decision has been taken, the user will be communicated the same along with a generated new token.
V. If the data presented to the model satisfies the required conditions, then the decision of service provision may be communicated the concerned banks.
VI. After receiving the information from the authentication server i.e. the initial bank server, the banks will start direct communication with the user for providing the services.

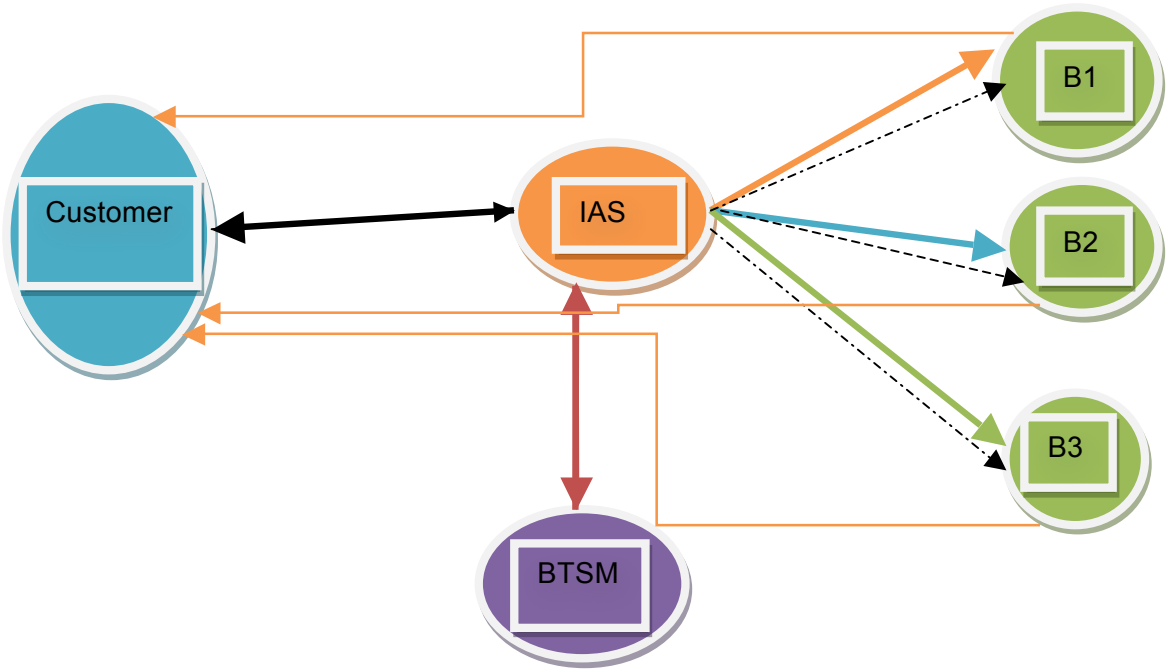


Figure 2: Proposed e-multi-banking model

Multiple banks offer their products/services to the customer with the help of the intermediary, which additionally acts as mediator between those entities by facilitating the multi-tasking process. Hence, the customer delegates to an authoritative IAS (Bank server) which proceeds with the distribution of client request messages to Bi and she is in charge of compounding the bank's responses as a unique message to C.

As shown in Figure 2, our model considers that banks Bi have no direct connection to BTSM (Banking transaction service model). Thus, all the payments authorization information should be transmitted to the transaction service processor through the client application. The client application starts the process after the customer checks out the services of various banks i.e., the list of banks available for service.

Advantages of the proposed model over the existing model:

1. Initial additional investment required minimized. With this model, it would be possible to upgrade the existing system by adding some important features.
2. As the bank server acts as an intermediary server, it can be trusted to a good extent.
3. The user interaction could be much easier.
4. User satisfaction would be a plus.
5. Risk of offline dictionary attacks could be reduced.

BTSM (BANKING TRANSACTION SERVICE MODEL) FOR AUTHENTICATION AND FRAUD DETECTION

BTSM (Banking Transaction Service Model) is used by the Intermediate authentication server. The purpose of this is authenticating a user and notifying the banks to provide or deny the services. The IAS initially gives the acquired data from the customer/user to the model. The IAS then identifies the feasibility of providing the services to the particular user or denying the services based on the decision given by the model.

The model used by IAS considers various parameters related to the user and also the associations as well as his relationships. The various aspects that the model would consider are:

a. The past transactions of the user with the banks.
b. The current balances maintained in each bank.
c. Past guarantees given by him/her.
d. Current guarantees given by him/her
e. Past commitments to be as security for someone else.
f. Current commitments to be as security for someone else.
g. The history of the customers to whom the guarantee has been given.
h. The history of the customers to whom the security has been given.
i. Past loans taken from the banks.
j. Current loans taken from the banks.
k. Loan repayment history of the user.
l. Loans taken by people to whom the guarantee/security has been given.
m. Loan Repayment history of the customers to whom the guarantee/security has been given.
n. The information about the users who are related to the user.
o. History of the user's family members with other banks.

The model which is used by the Intermediate Authentication Server could comprise the following:

a. Current Information of the customer/account holder
b. Past Information
c. Transactional Information
d. Account Information
e. Relationship with the customers of the same bank or the other banks.

As shown in the figure 3, the basic things which are listed above can be split further like : loan information, information of security and guarantee given to others currently as well as in the past. account information, transactional information with the other banks or the same bank and relationships with the customers of the same bank and also with those of other banks. The general representation of transactions performed by three users is shown in the table 1. Tables 2, 3, 4, 5 show the various possible combinations of the transactions performed by the user and the conditions.

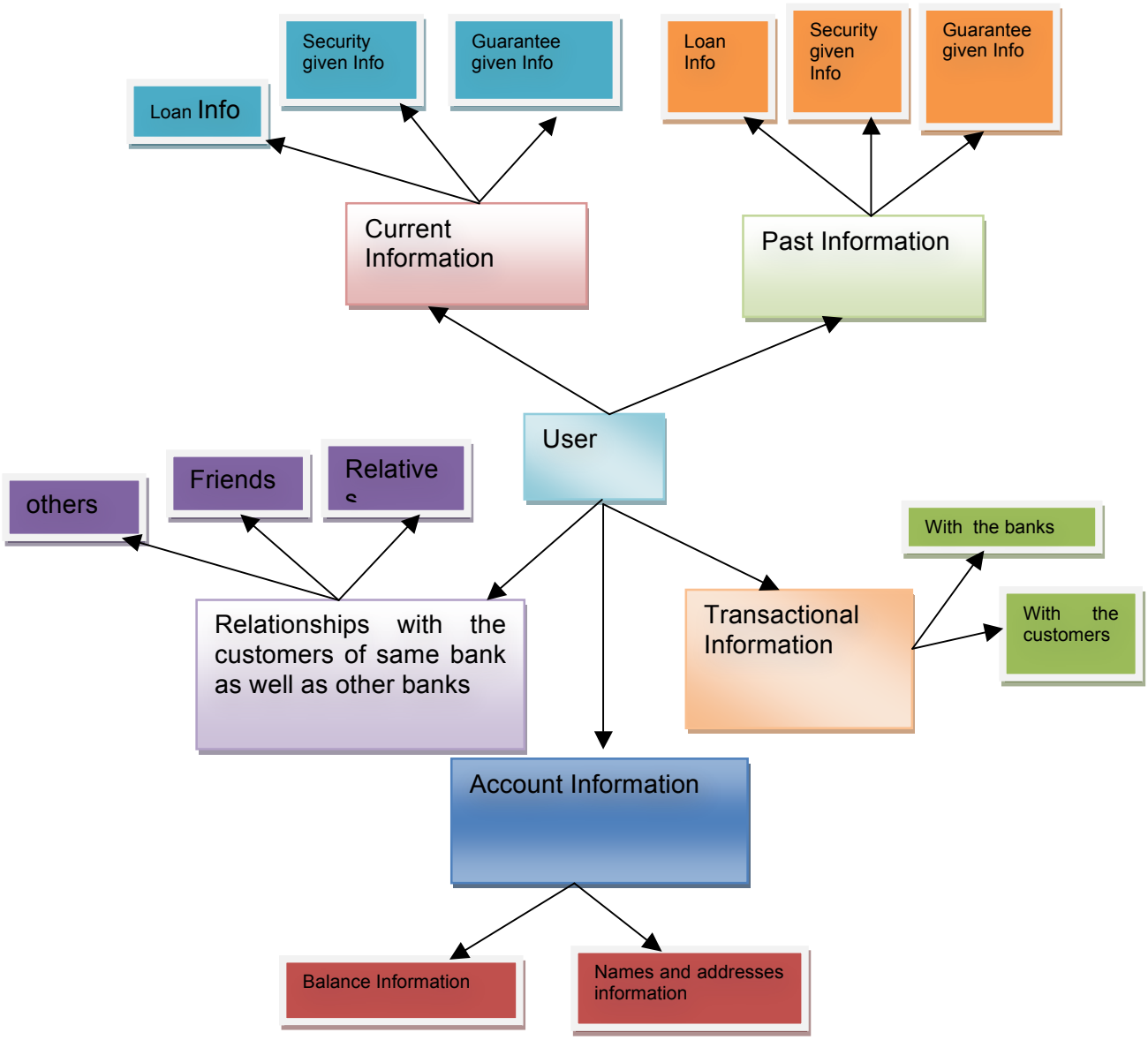


Figure 3: Representation of various parameters which are analyzed by the model

PROBABILITY OF SUCCESS IN ACCESSING THE EXISTING SYSTEM AND IN SECURING THE PROPOSED SYSTEM

Whenever a data is sent from a particular user to the authentication server, there is a probability for the authentication server to identify whether the user is genuine or fraud. The user behavior analysis is done. It is an inevitable fact that if more number of factors is involved in authentication, then the probability that the system is being attacked would be less. Analysis- based simulations are done using the lab view and the results are presented further in this paper.

The authentication server gets the data. If the trusted authentication server itself is compromised, then someone can try to hack the entire data. We calculate the probability of a fraudulent user being successful in an attempt by considering the total number of attempts made.

Let n be the sample space i.e. number of possible combinations that can be made by the user.

Let p_k is the probability of success in **k^{th}** trial.

Then

Probability of success in first attempt **$(p_1) = 1/n$**

$$\Longrightarrow \quad n=1/p_1 \qquad (1)$$

Probability of success in second attempt $(p_2) = 1/(n-1) = 1/(1/p_1-1) = p_1/(1-p_1)$

Probability of success in kth attempt **$(p_k) = 1/(n-k+1)$** (2)

This signifies that the probability of success slightly increases as the number of attempts keeps on increasing which appears to be quite obvious. The same is depicted in the figure 5a and 5b for n=100 and n=1000 respectively.

In figure 5a, as k⟶n the probability of success in kth attempt approaches exponentially to 1 which is quite obvious. Whereas in figure 5b the variation is almost linear as k is much less than n.

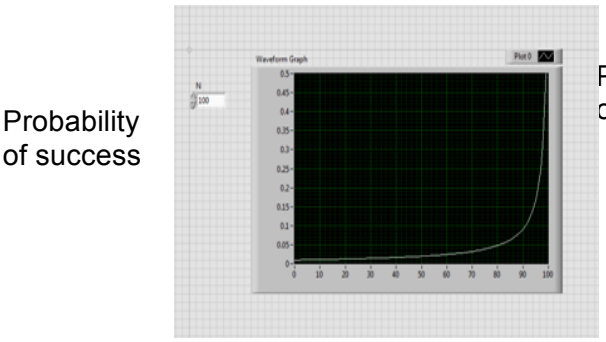


Figure 5a : Graph with 100 (n)samples

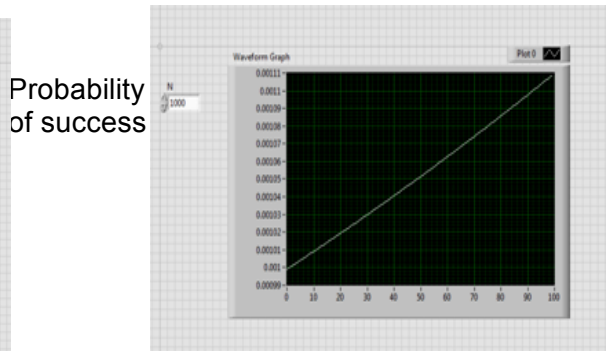


Figure 5b : Graph with 1000(n) samples

In the above discussion we computed the probability of success in the individual attempt. The logic can be extended for computing the probability of success resulted in a particular attempt such that all previous attempts are failed.

From the conditional probability, we have [13]

$$p(A|B) = p(A \cap B)/p(A) \quad \text{and} \quad p(B|A) = p(A \cap B)/p(A) \qquad (3)$$

From the above we have

$$p(A \cap B) = \begin{cases} p(A|B)p(B) \\ (B|A)p(A) \end{cases} \qquad (4)$$

Thus, we have $p(A|B) = (p(B|A).p(A))/p(B)$ (5)

Assuming,
A = Success in the kth attempt
B = Failure in (k-1)th attempt

Then
We have p(B|A) = 1 by assumption that success in A is possible only when B is failure.

Thus,
We have $p(A|B) = p(A)/p(B)$ from (5)

Let us assume probability of A = p(k), probability of B = 1-p(k-1)

Now we represent p(A|B) = p(k/(k-1))
then we have
$p(k/(k-1))=p_k/(1-p(k-1))$ (6)

$\Longrightarrow (1/(n-k+1))/((n-k+1)/(n-k+2))$ (7)
$= (n-k+2)/ (n-k+1)^2$

The above equation (7) will be used further to compute the probability of success in securing the system using the proposed technique i.e. the use of data mining model. So far we have computed the probability of successful attempt in breaking the security aspects using one of the n possible combinations. In the proposed system we add one more parameter which is dynamic in nature to provide or deny the services. The parameter that is considered here is the Data mining. Here, though the user could break the password, the services provision will be further verified based on the decision of the data mining model. This is likely to increase the security aspect of the system. We assume this data mining model is independent upon the existing single parameter passwords.

The data mining techniques further reduces the probability of the successful attempts by adding a factor c. Then we recalculate the above equation 3, 4 and the same is discussed below.

Let c be the factor to be introduced,
such that
0<=c<=1
From (7)
$p_k/p_{k-1}=(1/(n-k+1))/((n-k+1)/(n-k+2))$

We can calculate $p_k/p_{k-1} = (c/(n-k+1))/(1-(c/(n-k+2))$

This implies $\Rightarrow$ $(c(n-k+2))/((n-k+1).(n-k+2-c))$

The variation of the probability is shown in the figure 6 for different value of c. From the graph we can understand that, after adding one more parameter namely c, the probability of securing the system is increasing irrespective of the number of attempts made. The reason behind this is that the importance will be given for the user behavior, past data etc, by the BTSM model that is proposed. It is a proven fact that the parallel and distributed computing enhances the cyber security [14]. So, to achieve the best level of security, we should use the computational results that are parallel computed when user requests for the provision of services.

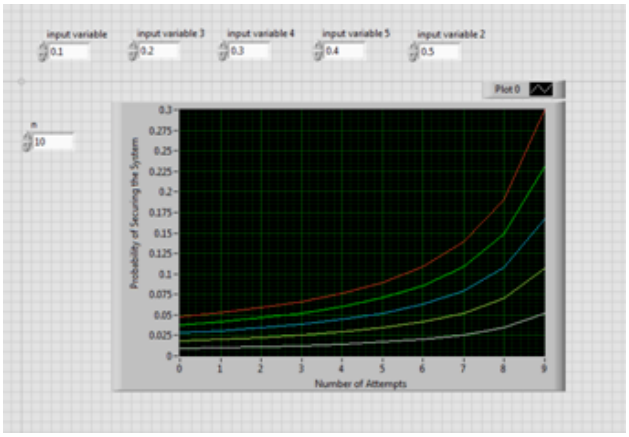


Figure 6: Graphical representation after adding parameter c

SIMULATION

A rich domain model is required to facilitate the decision making process and an expressive plan representation mechanism for fraud detection. The prime requirement of the model is the real time response to the unpredictable events. Petri nets are a powerful tool for modeling and analysis of discrete event systems. Petri nets allow a structured representation for static knowledge and in addition, they provide a systematic procedure for reasoning, which captures the dynamic behavior of the domain (here e-banking domain). The ability of Petri nets to represent causality, concurrency and conflict relations explicitly makes them an excellent tool for representing the planning problem or a decision making problem. The decision making process can be defined either by a series of actions that cause a system to achieve or maintain a goal state or by a

sequence of state transitions starting from an initial state to a goal state. In this work, the process is represented by a series of actions.

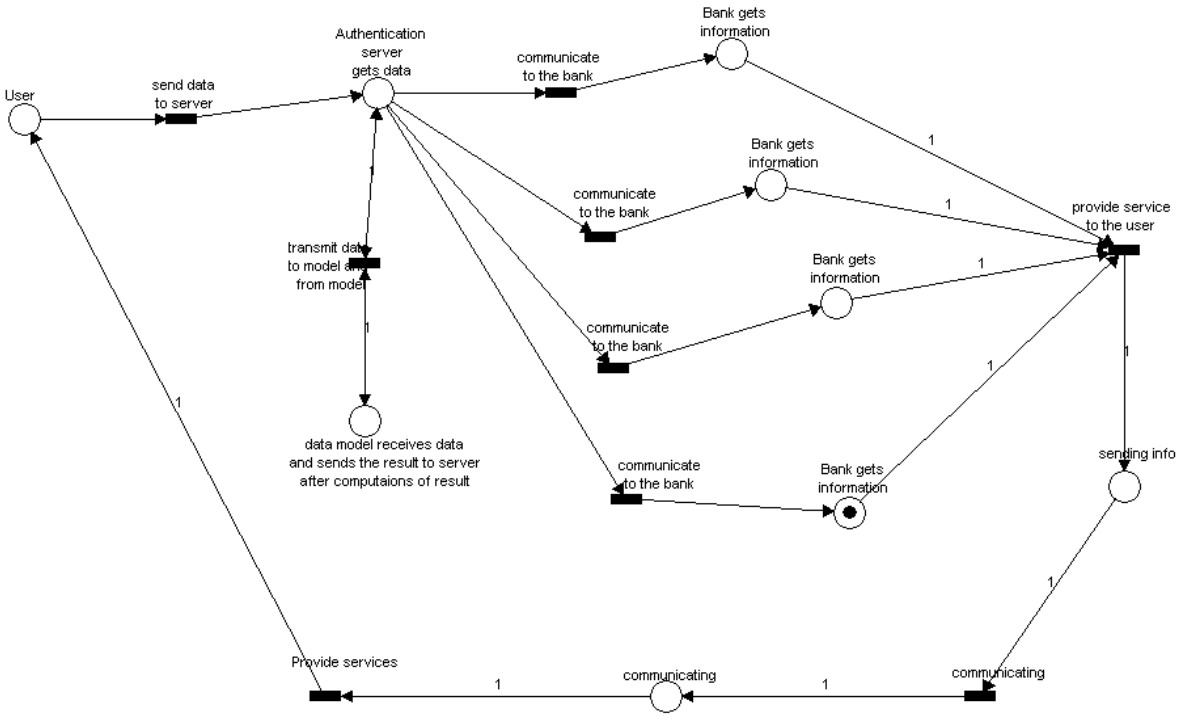
Places Used: Various places which are being used to represent different states are

a) User
b) Authentication server gets data
c) Data model receives data and sends back the result to the authentication server after necessary computations.
d) Banks get information
e) Sending Information
f) Communicating to the user.

Transitions Used: Various transitions used to represent different processes that are being carried out are:

a) Send data received from the user to the authentication server
b) Transmit the data to the model and from the model to the server
c) Communicate the decision to the bank if user is successfully authenticated
d) Provide or Deny the service to a particular user

The Simulations are done using the simulation tool HPSIM. During the process we observed that the flow of information is initially going to be from the user to the intermediate authentication server and then to the banks requested by the user.



CONCLUSION

We conclude by saying that electronic multibanking security could be enhanced with the help of trusted third party authentication server. The third party server can be any of the bank server's. Use of Data Mining techniques for decision making could improve the security aspect of the system. Multi-banking is an emerging area which is drawing the attention of both the user's as well as the service providers. In order to provide better services with a greater security banks have to go for novel concepts and models, obviously. Data mining models are one of such models that can be used effectively for fraud detection and identification of fraudulent activities. In future, we would like to enhance this work by analyzing various possible combinations of relationships among the customers and also how data mining can be effectively used in multi-banking, so that security can be at par with the level of services provided by the banks. We also would like to work on privacy preserved data mining for multi-banking.

ACKNOWLEDGEMENT

We would like to thank our college management for providing us an opportunity to carry out active research work in the area of multibanking. We also would like to thank our beloved principal Prof.K.N.B.Murthy and our HOD Dr.P.Punitha for their continuous and constant support. Last but not least, we wish to thank everyone who has co-operated with us.

REFERENCES

Lingling Wang, Coll. of Inf. Sci. & Technol., Qingdao Univ. of Sci. & Technol., Qingdao, China A new multi-banking e-cash protocol using anonymity control

Jiawei Han & Micheline Kamber, Data mining:concepts and techniques.

D. Foster, and R. Stine. "Variable Selection in Data Mining: Building a Predictive Model for Bankruptcy", Center for Financial Institutions Working Papers from Wharton School Center for Financial Institutions, University of Pennsylvania, 2002.

Mahmood H Shah Ashley Branganza Cranfield University, UK , Sajid Khan Cardiff University, UK, Mark Xu University of Portsmouth, UK , A Survey of Critical Success Factors in e-Banking.

Maher Aburrous, M.A. Hossain,Keshav Dahal, Department of Computing University of Bradford, Bradford, UK and Fadi Thabtah, MIS Department Philadelphia University Amman, Jordan, Intelligent phishing detection system for e-banking using fuzzy data mining, Elsevier Expert Systems with Applications 37 (2010) 7913–7921.

S.J. Hong, and S. Weiss. "Advances in Predictive Model Generation for Data Mining" , Proceedings 1st International Workshop Machine Learning and Data Mining in Pattern Recognition, 1999

B. Zupan, J. Demsar, M. Kattan, M. Ohori, M Ian H Witten & Eibe Frank, Data Mining-practical machine learning toold and techniques.

Peter Andriaans & Dolf Zantinge, Data mining.

Xiaozhe Wang, Ajith Abraham, Kate A. Smith, Intelligent web traffic mining and analysis, a Journal of Network and Computer Applications 28 (2005) 147–165.

Supriya Kumar De, P. Radha Krishna, Clustering web transactions using rough approximation, Fuzzy Sets and Systems 148 (2004) 131–138

Hui Xiong, Member, IEEE, Gaurav Pandey, Student Member, IEEE,Michael Steinbach, Member, IEEE Computer Society, and Vipin Kumar, Fellow, IEEE, Enhancing Data Analysis with Noise Removal, IEEE transactions on knowledge and data engineering, vol. 18, no. 3, march 2006.

John shafer, Rakesh Agrawal, Manish Mehta, SPRINT:A scalable classifier for data mining, Proceedings of the 22nd VLDB conference, Mumbai, India, 1996.

http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/Chapter4.

Vipin kumar, University of Minnesota, Parallel and distributed computing for cybersecurity, IEEE DS ONLINE EXCLUSIVE CONTENT, Invited article:Security.