A Novel Approach Using Advanced Modeling Traverse Tree For Controlling The Accuracy Level Of Traffic

[1] M. New Begin [2]R.Varalakshmi [3]X.Agnes Kala Rani
[1,2]Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi, Chennai, India.
[3]Karpagam university, Coimbatore, India
[1]newbegin_m@yahoo.com [2]varal90@gmail.com [3]agneskalarani@yahoo.co.in

Abstract : **In recent years the traffic in the urban areas is worsening. In order to increase the accuracy, quality, customer satisfaction over the pollution that was caused by the traffic is highly necessary. Previous work on multiple decision trees shows that it is impossible to some of the cases especially in the accuracy and the quality. We strongly exploiting the false fact of the previous concept to make a new technical conclave named as Advanced modeling traverse tree that is the combination of multiple random decision trees. The proposed concept can be possible in any other multiple setup of trees. The proposed system uses the six sigma methodology to make it as better and successful security solutions. Six sigma is a method of data management process that can be used to achieve the goal of near perfection in process performance. This paper mainly focuses about the improvement or the control over the traffic, in various specifications. Added that it improves the customer level satisfaction, non-polluted process cycle and the quality.**

Keywords: six sigma, accuracy, conclave.

I. Introduction

Data mining is increasingly popular because of the sizeable part it can make. It can be used to control costs as well as contribute to revenue increases[6].The actual data mining task is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns such as groups of data records[7]. Privacy preserving data mining has emerged as an effective method to solve this problem[1].In previous analysis the random decision tree was used to give partial model with medium cost. But the problem occurs with it is the customer do not have better satisfaction, and do not have congestion control. In this paper we propose a new technique Advanced modeling traverse tree to overcome the existing problems. In addition we propose the six sigma methodology to make it as better and successful security solutions. The use of Advanced modeling traverse tree is to control the traffic and cost. The traffic congestion problem in urban areas is worsening since traditional traffic signal control systems cannot provide efficient traffic control[6]. .Our procedure is to control the traffic of a large connected intersections and the result obtained is capable. The average delay time can be reduced by 42.76% compared to the predictable fixed structure traffic signal and 28.77% compared to the vehicle actuated traffic control approach. And it improves the better customer satisfaction and efficiency.

Six Sigma seeks to improve the quality of process outputs by identifying and removing the causes of defects (errors).It reduce to smaller cost[3].Our main influence is to provide good security with very high efficiency.

II. Advanced Modeling Traverse Tree

Advanced modeling traverse tree (AMTT) is the process of visiting the destination node within the tiny time. In AMTT each nodes contain a weightage. Based on the weight we can visit the node and we can reach the destination quickly without any congestion. In this concept we are using two trees and by linking those two trees we can reach the destination by travelling from one tree to another tree. In the following figure, its shows how the AMTT works. In the tree each nodes contain some weight. Based on those weights we can traverse tree swiftly using the shortest path by calculating the weights of the nodes that we have travelled. The description about the fig 1 as follows:From A->B->D->G the time it takes is 12.By using the AMTT it reduces the travelling time to 7 and also it reduces the traffic in the tree. The relative weight given to classifier accuracy whereas the average depth of the tree will reflect the weight given to efficiency[2].

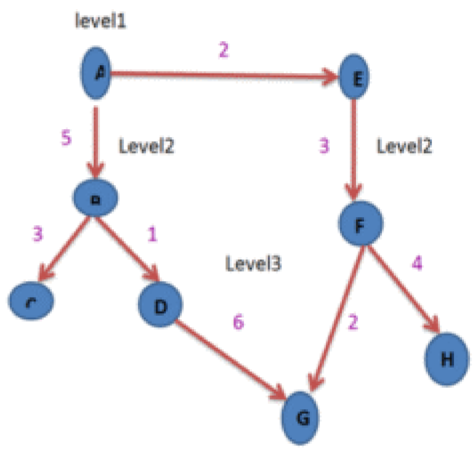

Fig1: Advance modeling traverse tree

III. Six Sigma

Six Sigma seeks to progress the quality of process outputs by identifying and removing the causes of defects [4]. The DMAIC project methodology has five

ISSN: 2278-0181

phases[11][5]. Define the system, the voice of the customer and their requirements, and the project goals, specifically. Evaluate key aspects of the current process and collect significant data. Analyze the data to explore and verify cause-and-effect re. Determine. Seek out root cause of the defect under exploration. Improve or optimize the current process based upon data analysis using techniques such as experiments or mistake proofing, and standard work to create a new, outlook state process. Control the future state process to ensure that any deviations from target are corrected before they result in defect. Implement control systems such as statistical process control creation boards, illustration workplaces, and continuously monitor the process.

A. *Advantages*

1. Maximize the customer satisfaction.

2. Maximize the product quality.

3. Save the cost.

4. Faster life cycle time.

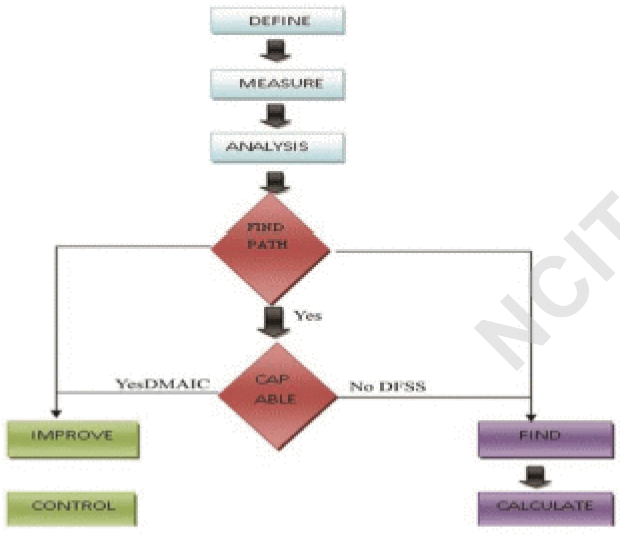

Fig 2.DMAIC Flow Diagram

DMAIC refers to improving, optimizing and stabilizing the tree processes and designs. DMAIC is an tightening of the five upgrading steps: Define, Measure, Analyze, Improve and Control.

The description about the fig 2 as follows. All of the DMAIC process steps are vital and always proceed in this order[6][3].

1. DEFINE: Defines the overall tree formation for the tree traversal.

2. MEASURE: Measures the weightage of each and every node of the tree that is to be traversed.

3. ANALYSIS: Analysis is made after travelling each node by calculating the weight caused after travelling the tree.

4. IMPROVE: Identify resourceful solutions to fix and avoid process problems. That is improving the enhanced traversal by dipping the total time caused to travel without the traffic.

5. CONTROL: Traffic control is made and cost is controlled efficiently.

IV. PROBLEM STATEMENT

Since the RDT code can be used for multiple data mining tasks, we focus on classification for ease of discussion. The basic problem in distributed classification is to train a classifier from the distributed data and then classify each new node. In the privacy-preserving case, the additional constraint is that the process of building the classifier, or of classifying an instance should not leak any additional information beyond what is learned from the result (and the local input).Clearly, the specific steps are dependent on the ways in which the data is distributed. The two most common are horizontal and vertical partitions. When data is horizontally partitioned between k parties, each party holds different instances, but collects the same pieces of information. All parties share the schema, though the specific transactions in their local databases are unique. In this case, we assume that there are no overlapping transactions. Clearly, since the schema is shared by all parties, the class attribute C is also known to all parties. Design a decision tree that correctly classifies all the training samples, also known as a perfect tree[4].

A. *Horizontally partitioned data*

Here, each party knows the structure for all of the trees. Indeed, if a party thinks that a tree reveals too much information, it might wish to reject that particular tree and ask for alternatives. This can be achieved using an electronic voting protocol[5] if it is necessary to ensure that no participant learns which other participant(s) rejected any tree. Note, however, that rejecting a tree itself leaks information. Even if it is not known who rejected the tree, any party who does not have an issue with the tree knows that at least one other party was not satisfied with the tree. In the worst case, if there are only 2 parties, one party will know that the tree causes privacy breach for the other party. However, it is not known which path of the tree may lead to privacy leakage. In any case, this is unavoidable since a party can either accept a tree or reject it (and accepting in such cases may lead to larger leakage).Once all parties agree on all of the trees, the computation of the leaf values proceeds. To do this, all parties locally compute the values for the leaf nodes. Following this, they simply have to add up all of the leaf values to get the global tree values. This can be easily done in a secure fashion using the Secure Sum protocol[8].Since all parties have the information for all of the trees, classification is simple and can be locally performed.

ISSN: 2278-0181

B. Vertically Partitioned Data

With vertically partitioned data, all parties collect data for the same set of entities. However, each party collects data for a different set of attributes[9]. Now the parties cannot independently create even the structure of a random tree, unless they share the attribute in sequence among each other. Thus, there are two possibilities reason:

• All parties share basic attribute information (i.e., metadata). Now they can independently create random trees.

• There is no sharing of information. Now, the parties need to collaborate to create the random trees. These trees could themselves exist in a distributed form. Unlike the horizontal partitioning case, the structure of the tree does reveal potentially sensitive information, since the parties do not know what the attributes are owned by the other parties. Therefore, we directly address the case of fully distributed trees.

V. HYBRID HORIZONTAL VERTICAL ALGORITHM

By using the hybrid horizontal vertical algorithm we can find the routing path quickly. The following algorithm shows how it will work[10].

Step1: M and N are two nodes with respect to horizontal and vertical tree.
Step2: Find the number of nodes coming from horizontal tree.
Step3: Find the number of nodes coming from vertical tree.
Step4: Calculate the average number of nodes arrived at routing. $M=\sum_{i=1}^{n}((v1+v2+v3+\cdots vn)\div vn)+\sum_{i=1}^{n}((h1+h2+h3+\cdots hn)\div hm)$
$M=\sum_{i=1}^{n}[((v1+h1)(v2+h2)\ldots(vn+hm))\div vn+hm]$
Step5: Searching the neighboring nodes n1,n2,n3,n4….nn
Step6: Measuring the cost of every path with dynamic allocation c1,c2,c3…cn
Step7: Based on the case average time has to take.
Step8: if(nodes>n)then
Average time=(n1+c1)(n2+c2)+……+(nn+cn)*60)+m
Step9: End if
Step10: Close the process

VI. RESULT AND DISCUSSION

There has been extensive work in privacy preserving data mining. Several data mining algorithm have since been proposed[1]. Recently, there has been a renewed interest in this field, a good discussion can be found[4]. The figure shows the comparison of various tree. Based on weightage it travels.

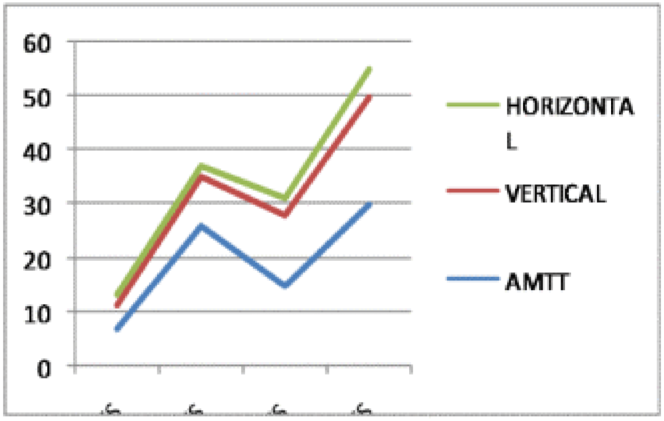

Fig: 3 Comparison of shortest path in various trees

A. Cost Table

The cost table shows the difference between the random decision tree and advanced modeling traverse tree. For 7nodes random decision tree cost is 9 and for same nodes the advanced modeling traverse tree shows the cost 4.

Table 1: Cost table

S.no	Nodes	RDT	AMTT
1	7 nodes	9	4
2	10 nodes	9	6
3	15 nodes	10	5

[Here RDT means Random Decision Tree and AMTT means Advanced Modeling Traverse Tree]

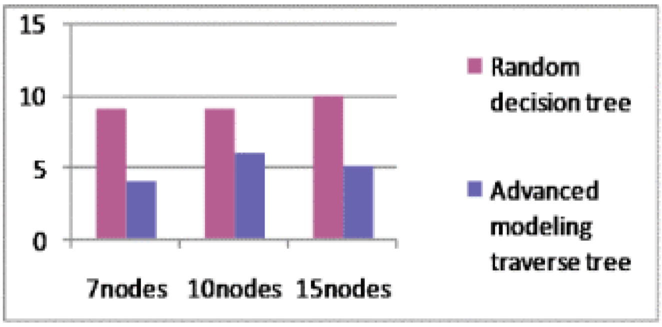

Fig 4: Comparison of cost in various tree

The figure 4 shows cost variation for random decision tree and advanced modeling traverse tree.

ISSN: 2278-0181

B. Accuracy Table

The accuracy table shows accuracy level of two trees and its shown in the line graph. The blue colored line for random decision tree and brown colored line for advanced modeling traverse tree. By using the advanced modeling traverse tree accuracy is more.

Table 2: Accuracy table

S.no	Nodes	RDT	AMTT
1	7nodes	2	4
2	10nodes	3	7
3	15nodes	4	9

[Here RDT means Random Decision tree and AMTT means Advanced Modeling Traverse Tree. The following graph shows accuracy levels of trees]

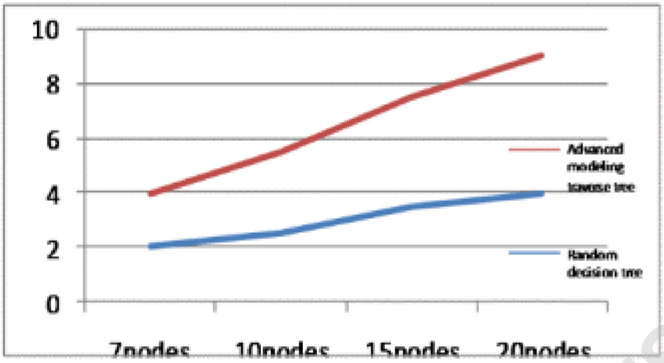

Fig5: Comparison of accuracy in various tree

VII. CONCLUSION

We have demonstrated that general and efficient and distributed advanced modeling traverse tree discovery is truly feasible. We have considered the security, cost, traffic when dealing with trees. Our approach leverages the fact that AMTT in structure provide congestion control with less computation. The experiments show that the AMTT algorithm requires less cost, congestion free. In the future, we plan to develop general solutions that can work with arbitrarily partitioned data and also highly confidentially data.

REFERENCES

[1] J.Vaidya, c.Clifton, and M.zhu, privacy –Preserving Data Mining, 1sted, Ser.Advances in Information Security.Springer-Verlag, vol.19, pp:2005.

[2] P.Swain and H.Hauska,"The decision tree classifier design and potential,"IEEE Trans.Geosci.Electron, vol.1,pp:142-147,1977

[3] Li, M.-H.C. ,Feng Chia Univ., Taichung Al-Refaie, A. Cheng-Yu Yang , "DMAIC Approach to Improve the Capability of SMT Solder Printing Process" ,vol 1, pp126-133,2008.

[4] W.Du and M.J.Atallah,"Secure multi-party computation problems and their applications: A review and open problems, "in New Security Paradigms Workshop.NewYork, NY: ACM, vol 1, pp:11-20, September 11-132001.

[5] D.Gritzalis, Secure Electronic voting 1sted. Ser.Advances in Information Security.Springer-verlag, vol.7, 2003.

[6] http://www.dmaictools.com

[7] http://en.wikipedia.org/wiki/Data_mining

[8] B.Schneier, Applied cryptography, 2nded.Hoboken, NJ: John Wiley & sons,1995.

[9] http://www.en.wikipedia.org/wiki/verticalparitioneddata

[10] http://www.en.wikipedia.org/wiki/Algorithm

[11] http://www.esoonline.com/article/221094/ideas-you-can-steal-from-six -sigma