A Survey on DETECT – Discovering and Evaluating Trust using Efficient Clustering Technique for MANETs

K.Sudharson[1] Assistant Professor, and N.Partheeban[2] Associate Professor.

[1,2]Department of Information Technology, S.A.Engineering College, Chennai-77,India.
[1]mail@sudharson.in, [2]knparthi78@gmail.com

ABSTRACT

Analyzing and predicting behavior of node can lead to more secure and more appropriate defense mechanism for attackers in the Mobile Adhoc Network. In this work, models for dynamic recommendation based on fuzzy clustering techniques, applicable to nodes that are currently participate in the transmission of Adhoc Network. The approach focuses on both aspects of MANET mining and behavioral mining. Applying fuzzy clustering and mining techniques, the model infers the node's preferences from transmission logs. The fuzzy clustering approach, in this study, provides the possibility of capturing the uncertainty among node's behaviors. The results shown are promising and proved that integrating fuzzy approach provide us with more interesting and useful patterns which consequently making the recommender system more functional and robust.

KEYWORDS

MANET Mining, Fuzzy Clustering, Knowledge Discovery.

1. INTRODUCTION

A mobile ad hoc network (MANET), is a self-configuring infra structure less network of mobile devices connected by wireless links. Nodes in mobile ad-hoc network are free to move and organize themselves in an arbitrary fashion. Each user is free to roam about while communication with others. The path between each pair of the users may have multiple links and the radio between them can be heterogeneous. Web personalization is a process in which Web information space adapts with user's interests [7]. Same way in MANET the nodes are continuously participating in the routing is recorded in transmission log file with node's interest rate. Web usage mining techniques are widely employed for extracting knowledge about user interests [6]. Likewise, MANET mining Techniques are widely employed for extracting knowledge about the node interest. However, as the vagueness and imprecision in node interests are key features of MANET systems, traditional models which use hard computing techniques are inadequate. Since, the navigational behavior of nodes on the MANET is uncertain, using fuzzy clustering algorithms are more suitable for behavioral mining.

The rapid growth of communication networks such as the Internet and wireless mesh networks has spurred the development of numerous collaborative applications. Reputation and trust play a pivotal role in such applications by enabling multiple parties to establish relationships that achieve mutual benefit. In general, reputation is the opinion of the public towards a person, a group of people, an organization, or a resource. In the context of collaborative applications such as peer-to-peer systems, reputation represents the opinions nodes in the system have about their peers and peer-provided resources. Reputation allows parties to build trust, or the degree to which one party has confidence in another within the context of a given purpose or decision. By harnessing the community knowledge in the form of feedback, reputation based trust systems

help participants decide who to trust, encourage trustworthy behavior, and deters dishonest participation by providing a means through which reputation and ultimately trust can be quantified and disseminated. Without such mechanisms, opportunism can erode the foundations of these collaborative applications and lead to peer mistrust and eventual system failure. A rich variety of environments and applications has motivated research in reputation systems. The success of a reputation system is measured by how accurately the calculated reputations predict the quality of future interactions. This is difficult to achieve in an environment where any party can attempt to exploit the system to its own benefit. Some attacks have a narrow focus and only affect the reputation of the misbehaving identity or a few selected targets. This paper concerned with distributed association rule mining algorithm in finding frequent malicious behaviour of the node. The work is different from many existing algorithms where most of existing algorithms focusing only on the reputation-based mechanisms that improve security in MANETS. Our distributed algorithm focuses on association between the nodes in order to reduce the malicious nodes in Adhoc Network. Also we propose Clustering Techniques to identify the good nodes and grouping them accordingly and this allows us to recognize crucial spots in changing node behaviour time to time. This paper is the first survey focusing on the characterization of reputation systems based on mining techniques and threats facing them from a computer science perspective. Our work contributes to understanding which reputation system design components are vulnerable, what are the most appropriate defense mechanisms, and how these defense mechanisms can be integrated into existing or future reputation systems to make them resilient to attacks.

This paper is organized as follows .In section2; we discuss the background study the related works. In section 3 we discuss the proposed work explained to find the Node Behavior and finally clustering techniques also described in detail. Section 4 Discussion, Section 5 is Conclusion and future work.

2. BACKGROUND STUDY

In the recent years many researches tries to predict the node behavior. Castellano et al. have presented an approach for expressing similarity between Web users according to their access patterns. They considered the access time to a web page as a parameter in the model [4]. Castellano et al. have also used a neuro-fuzzy model to develop a Website personalization framework for dynamic suggestion of URLs to users [5]. Likewise, in this model it supposes to find the similarity between the nodes according to their accessing path. This approach considered the node access time on particular path as parameters to find the misbehaving node which are deviating from the normal routing time and mechanism. A different approach is purposed in the work of Kraft et al. in which they extracted fuzzy rules from user profiles and used them for information retrieval. Web user profiles usually reflect only access behavior of users and not keywords on a website. In this work, they rated web pages for building user interest profiles [3]. In Our approach the knowledge is extracted from the node profiles and used them for finding node behavior. A bi clustering approach to correlate web users and pages are purpose in the work of [9], their purposed approach is a three step process which concentrates on the principles of spectral clustering analysis and provides a fuzzy relational scheme for both user and page clusters but they didn't consider to active users for providing dynamic recommendation. Another approach for combination of content mining and web usage mining in order to predicting user navigational behavior is represented by [8]. The frequent word sequences are used for capturing textual content of web pages. Log files data are combined with content clusters for extracting rules about user's behavior. In our approach we concentrate on clustering analysis and provide fuzzy relation scheme for node and transmission path usage mining, Another approach for combination of path mining and path usage mining in order to predicting user navigational behavior. Log files data are combined with node clusters for extracting rules about node's behavior .Their work is improved in this work by employing fuzzy

techniques for node clustering and predicting rated recommendations for active node. In this paper we propose a model applying both node and path clustering using fuzzy techniques for better dynamic recommendation process. The following formatting rules must be followed strictly. This (.doc) document may be used as a template for papers prepared using Microsoft Word. Papers not conforming to these requirements may not be published in the conference proceedings.

3. PROPOSED MODEL

The ADOHC personalization model is to provide the paths transparently provides node with the personalized environment. When node requests a path, the neighbor node retrieves and replies the requested path followed by a list of paths which may be interested to the node.Two major steps of a ADOHC personalization system are knowledge discovery and recommendation. In first step, node preferences are identified using log data called path usage data (PUD).In the next step, the achieved knowledge is used to identify the possibly interested paths to the nodes. This recommendation can be done in different ways such as adding related hyperlinks to the last web page requested by the user [2]. In this paper, Represented method dynamically recommends the highest match score paths to active node. Represented method steps are described in the following sections:
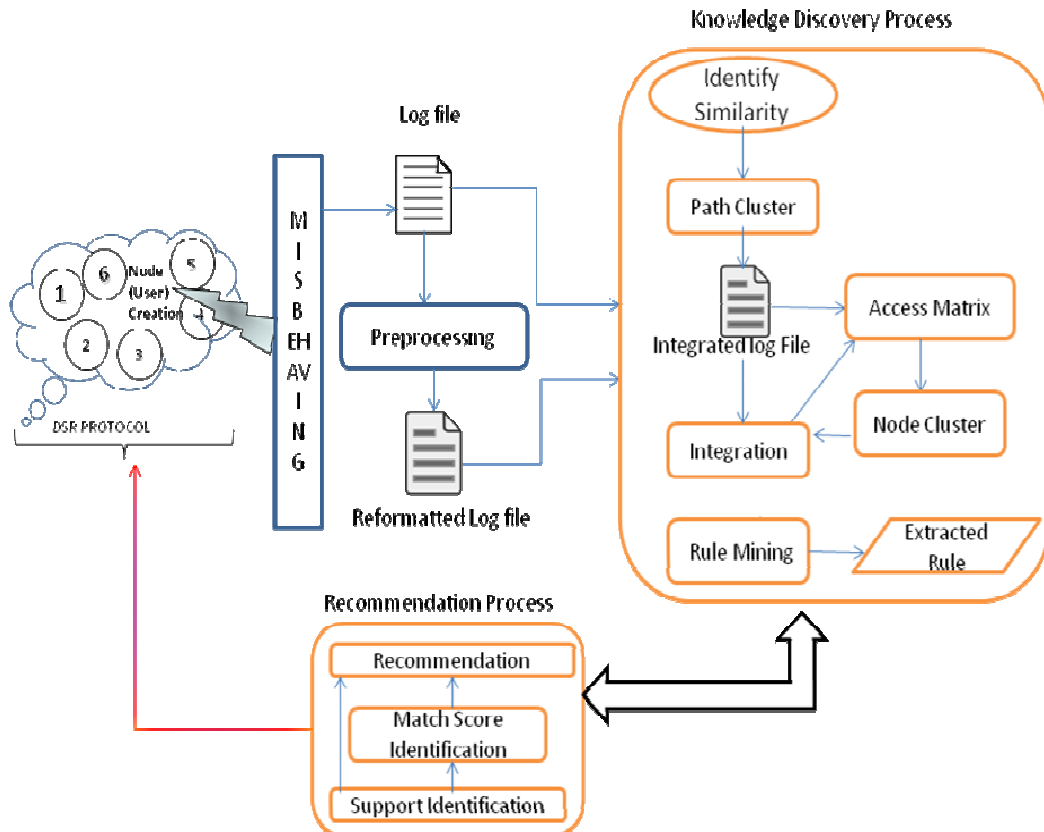
3.1. System Architecture

Fig a) Discovering Misbehaving Node Using Knowledge Extraction and Recommendation Process

3.2. Knowledge Discovery Process

In this step, using the previous behavior of each node, knowledge that represents node's preferences is extracted. For extracting knowledge from path usage data, 6 steps will be performed as described in the following:

3.2.1. Preprocessing

Log files have useful information about access of all nodes to specific routes. Extracting these information, reformatted log file which contains useful information such as "time, date, accessed path and IP address" is formed and useless requests such as accesses to different paths are removed from log file in data cleaning process.

3.2.2. Path Clustering

In this work, we used mining approach for path clustering. Assume $P = [p1, p2, ... pk\}$ is the set of k routes. Applying 3 following steps, paths were grouped in path based clusters.

Stepl. Assign each path to a single node.

Step2. Merge primary clusters based on the Jaccard coefficient similarity measure. Defined as:

$$sim(p_x, p_y) = \frac{|p_x \cap p_y|}{|p_x \cup p_y|}$$

$|px \cap py|$ represents the number of common path and $|px \cup py|$ represents total number of paths between two basic nodes.

Step3. The second step repeated until all paths being clustered into a pre defined number of path clusters. $DC = \{DC_1, DC_2, ... DC_n\}$ is the result set. Each $DC|$ represents a set of paths with similar distance.

3.2.3. Integration

In this step, the previously obtained path clusters are merged with reformatted log file and according to the result, access table can be produced. Then, nodes are clustered based on their behavior in access to path clusters. Assume U= {UI, U2 , ... , Um} as a set of nodes where each node is identified by distinct IP address and A= { A1, A2, ... , Am } is a set of node's accesses to path clusters. For example A_k indicates the access list of node, *k,* to a subset of path clusters $(A_k \varepsilon DC$, for K=1 to n). Access table is represented as a matrix with **m** rows and **n** columns where each entry represents the interest degree of *i-th* node to *j-th* path cluster. We use the interest degree which proposed by Castellano et aI., They defined it as the ratio of the number of accesses to each path cluster to the total number of accesses to all path clusters for each node [6].

$$Val(DC_i, U_i) = \frac{|\{DC_j | DC_j \in A_i\}|}{|A_i|}$$

Calculated matrix will be the input data for clustering algorithm.

3.2.4. Node Clustering

According to access matrix, nodes with similar interests can be clustered together. In this work, we used fuzzy c-means (FCM) method for clustering nodes. UC= {UC$_1$,UC$_2$,...,UCn} is result set where each **UCj** represents a set of nodes with similar interesting patterns.

3.2.5. Final Preparation

In reformatted log file, for each log entry, accessed path were substituted by most related path cluster and similarly, IP address will be replaced by the most relevant node cluster. The result of this step is input data for rule mining step.

3.2.6. Rule Mining

Varity number of algorithms such as Apriori, Eclat and FP-Growth are proposed by researchers for generating association rules [1]. Extracted rules show that which group of nodes and when are interested to what kinds of path clusters. We used apriori algorithm for mining our frequent item set.

3.3. Recommendation Process

Using results obtained from previous step, some related information can be recommended to a group of nodes. For example, {Node Cluster=UC7} => {path Cluster=DCl} means that paths in path cluster 1 can be recommended to nodes in node cluster 7.

3.3.1. Support Identification

If a new node starts a transaction, our model matches the new node with the most similar node clusters and provides suitable recommendations to him. Thakur et al. represented a criterion called support value which was calculated through following steps: [2]

Step 1: Assign active node to a new cluster. **UC$_{new}$ = {Unew}.**

Step2. Complete access matrix for active node according to the equation 2.

Step3. Calculate support value of UCnew to existing node clusters **UC$_i$** using Equation 3.

$$support(UC_{new}, UC_i) = \frac{\sum |user(val(DC, UC_{new}) - val(DC, UC_i))|}{UC_i \cup UC_{new}}$$

val(DC, UQ) shows the interesting value of nodes in **UC$_i$** to the documents in selected DC.

3.3.2 Match Score Identification

The match score calculation defines highest match node cluster for active node. Match score criterion is represented by Thakur et al. This parameter between **UC$_{new}$** and **UQ** is defined as following [2].

$$match(UC_{new}, UC_i) = 1 - support(UC_{new}, UC_i)$$

This give us a list of corresponding node clusters from the highest match score down to lowest match score.

3.3. Experimental Discussion

The work of taherizadeh et al. has improved in this paper by employing fuzzy clustering and supporting active node by a range of rated relevant recommendations. The results of this path based recommender system have been successfully modeled in this paper. Suggested recommendation model combines path and node clustering using fuzzy approach. Applying model on the access log file of nodes, a set of rated path was derived from node preferences for dynamic recommendation.

The advantages of the proposed model are summarized as following. Firstly, this model acts effectively in identifying node preferences. Uncertainty among node interests is an important issue which is clearly considered in this approach. Our model captures the uncertainty in node navigational patterns and establishes an effective recommendation strategy. Additionally, this approach provides dynamic user clustering for active nodes and supports them with personalized environment. Moreover, any changes in path can be simply informed to all group nodes which are highly interested on that particular path. While the advantages of the model are clear, there are some limitations in this research area that we consider them as future research directions. Path clustering and node clustering in this approach are considered in separate processes which is time consuming. In order to solve this problem a Two-way clustering method may be used.

This strategy in the same time not only clusters the objects but also the features of the objects will be clustered. Next common limitation in clustering algorithms is defining a suitable number of clusters which can be surmounted by employing QT clustering algorithm which does not require specifying the number of clusters and when running several times, always returned result is as the same.

4. CONCLUSIONS

In this paper, a model for dynamic recommendation based on fuzzy clustering techniques, applicable to nodes that are in highly mobile and sparse environment was proposed. In this work, node mining was integrated to path usage mining for finding nodes interesting rules. In this way, we investigated fuzzy clustering techniques for node clustering based on the knowledge obtained from node's previous behaviors. Then a fuzzy technique was used for grouping a new node in an existing node cluster. Finally based on extracted rules the Suggested links was added to requested page and represented to current node dynamically. Our work will be extended by employing a neuro fuzzy strategy.

ACKNOWLEDGEMENTS

REFERENCES

1. B.Goethals, "Survey on Frequent Pattern Mining ", P.O. box 26, FIN-00014 Helsinki Finland, 2003.
2. B.K Thakur, S.Q Abbas, A.K Trivedi, "A Recommender System to Personalize the Environment of Web User", IEEE International Advance Computing Conference Patiala, India, 2009
3. D.H.Kraft, $J.$ Chen, M. Bautista, MJ., and MA Vila, "Textual Information Retrieval with User Profiles Using Fuzzy Clustering and," Intelligent Exploration of the Web, Heidelberg, Germany: Physica-Verlag, 2002.

4. G. Castellano, A. M. Fanelli, C. Mencar and M. Alessandra Torsello, "Similarity-based Fuzzy clustering for user profiling", IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology -Workshops, 2007.

5. G. Castellano, AM. Fanelli, P. Plantamura, M.A Torsello, A Neuro-Fuzzy Strategy for Web Personalization, Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, 2008

6. M. P. Singh, "Web Usage Mining and Personalization" , a chapter in Practical Handbook of Internet Computing, Munindar P. Singh (ed.), CRC Press, 2005.

7. O. Nasraoui, "World Wide Web Personalization", in J. Wang (ed), Encyclopedia of Data Mining and Data Warehousing, Idea Group, 2005.

8. S. Taherizadeh and N. Moghadam, 2009. Integrating web content mining into web usage mining for finding patterns and predicting users behaviors, Internationa/ Journal of Information Science and Management, Vol, 7, No.1, pp.51-66

9. VA Koutsonikola and A.I. akali, 2009. A fuzzy bi clustering approach to correlate web users nad web pages, IntJ.Knowledge and Web Intelligence, Vol.1, Nos.1/2, pp. 323

Author Biography

Sudharson Kumar received a Master Degree in Computer Science and Engineering from S.A.Engineering College, Anna University, Chennai. He is working as Assistant Professor in Department of Information Technology, S.A.Engineering College, Chennai. He has awarded Best Outgoing Student Award while doing his Master Degree, and he started his research in Adhoc Network while pursuing his Masters Degree and he presented papers in various National and International Conferences. He is member of various societies like IACSIT, IAENG etc., He applied various journals and waiting for publication.