iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VI **Month of publication:** June 2017

DOI:

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

A Deeper Insight towards Dynamic Network Analysis

Ruby K. Rajan[1], Salitha M. K[2]
[1]PG Scholar Computer Science and Engineering, Kerala Technological University
2Asst. Professsor Computer Science and Engineering, Kerala Technological University

Abstract*: Dynamic Social Network Analysis is used to find the organization and behaviour of dynamic social networks. It is multi-mode, multi-plex, huge and dynamic. It combines the ideas from different areas like Social Network Analysis (SNA), Link analysis (LA), Multi Agent System (MAS), etc. The nature of the nodes are changing and the tie between the nodes are also evolving. Dynamic Network Analysis find its application in predicting the future of existing networks, public health, social media analysis, etc. Various algorithms are developed for the analysis of these networks. This paper deals with the major advancements in Dynamic Social Network Analysis.
Keywords— Dynamic Social Network Analysis , Community, Temporal Network Analysis, Latent Space Model, Inference Attack.

I. INTRODUCTION

The huge availability of data made people to think about the chances of converting it into useful information. This gave rise to the field of data mining. Data Mining is the process of gaining knowledge from data through a series of steps which include data pre-processing, pattern evaluation, knowledge presentation etc. Data mining extracts interesting patterns and knowledge from large datasets. The pattern may be a single event or combination of events whose occurrence is different from the expectation. Data mining uses the available data to analyze it in different views. Then it uses the description of interesting pattern to find similar patterns from the database. It develops a prediction model from the historical information. The data that are taken for processing may be from databases, data warehouses, web pages or other sources. Data mining has numerous subareas like text mining, web mining, pattern mining, sequential pattern mining etc. The major application of data mining includes health care, market basket analysis, education, fraud detection, finance, criminology, bioinformatics etc.

Dynamic Social Network analysis makes use of web mining concepts. The network is structured using nodes and links/edges that connects the nodes. These network has the capability to simultaneously analyze multiple networks[11]. Figure 1 illustrates the structure of a Dynamic Social Network, in which the nodes represents individuals and the link represents the relationship between them. The links are probabilistic. The characteristics of the nodes changes. The existing links may disappear and new links may appear. One example of dynamic social network is the friendship network, where friendship between individuals may change over time. The order in which the ties occur also differs. The communities and nodes in the network are always evolving. The changes in the flow of one sub-network may affect other network related to it. The summarization process in dynamic networks is not an easy task. Also large amounts of data need to be analysed. Because of these challenges the methods used in traditional Social Network Analysis is not applicable in dynamic networks. Some of the algorithms, in dynamic social network analysis that addresses the challenges are described in this paper.

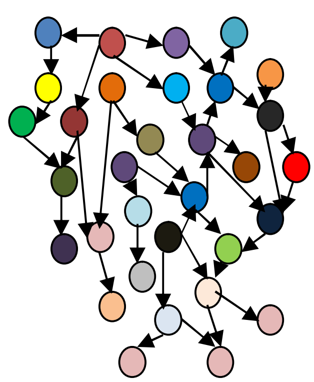


Fig. 1 Dynamic Network

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

II. COMMUNITY DISCOVERY IN DYNAMIC NETWORKS

Wang et al. [1] proposed a method, "NEIWalk" (NEI network based random walk) for finding community in dynamic networks. The method combines the structure of the link, the content of nodes and edges in dynamic content-based networks (social networks in which content information is present). A diverse random walk is performed in the transition probability matrix that represents the semantic upshot of the different type of edges in the network. The method first transforms the content-based network into NEI (Node-Edge Interaction) network. Figure 2 shows the transformation of content-based network into NEI network. NEI is a multi-mode network consisting of two node types, which are the nodes representing the node set and the nodes representing the edge set and three edge types. The edge types include those that are present in the original network, edges based on the similarity of the content of nodes and edges based on the similarity of the content of edges. The node content and edge content are represented using an n-dimensional feature vector and their similarity is measured by the cosine similarity between corresponding feature vectors.

The transition probability matrix is constructed to perform random walk in NEI network, which includes node content hop, edge content hop, structural hop from nodes to edges and edges to nodes. Since it is a dynamic network, the NEI network is updated according to evolving nodes and also the transition probability matrix. The community is discovered in two timestamps. In the first time stamp, the nodes are clustered on the basis of similarity measure calculated from the hitting time between the nodes. In the second timestamp, cluster label is obtained for each node by performing random walks. The node sets are partitioned to various clusters that are represented by the central node of all other nodes in the cluster. The labels for the clusters are generated by the k-center clustering method.

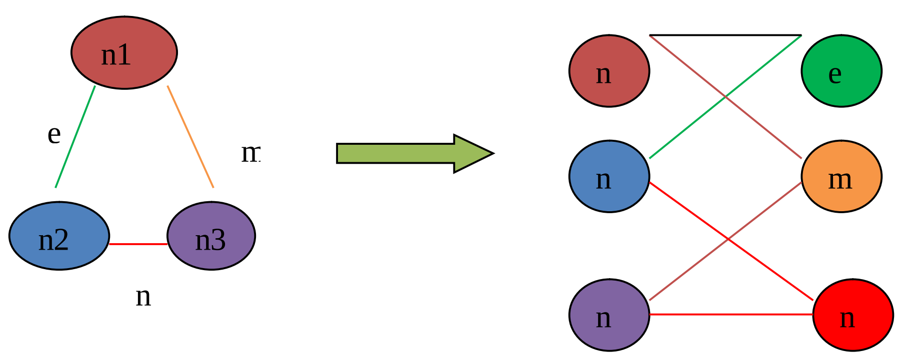


Fig. 1 Transformation of Content-Based Network

Tang et al. [2] proposed temporally regularized multimode clustering algorithm to address the problem of identifying the groups/communities that are evolving in dynamic networks. The algorithm initially generate a cluster indicator matrix, which is updated by a process that corresponds to Latent Semantic Analysis (LSA). The community structure of a mode at a particular timestamp is obtained using the community indicators of the modes that are interacting with it and the neighbouring timestamps. These constitute to the features of the nodes. LSA obtains the features of other modes and timestamp in a multimode network. The network updation is done using these details. The LSA is iteratively performed to reach an equilibrium state. This algorithm tend to produce a local optimal solution, by finding more accurate evolving community structures. This algorithm has scalability issue, when dealing with large-scale networks. The regularization is sometimes unable to detect unusual happening, which cause drastic changes.

Liu et al. [3] proposed evolutionary multi-mode clustering algorithm that makes use of the temporal details to detect the evolving communities in dynamic network. The algorithm provides a spectral clustering framework. The cluster indicator matrix is taken as continuous, so it is not constrained. The weighted features of clustering a particular mode is the clustering result of proximate timestamps and the related objects. This algorithm overcomes the scalability issue. But the framework is not able to execute the number of communities, the weights of each interaction and temporal details by its own. The user has to give such details.

III. SUMMARIZATION OF DYNAMIC NETWORKS

Qiang *et al* [4] proposed OSNet , an online summarization framework. The framework includes efficient algorithms for the summarization of large dynamic networks by considering only the high utility nodes or edges .OSNet includes online and incremental summarization algorithms for dynamic network using a tree model. A set of indexed spreading trees are used to model the interactions. The interestingness of the node is calculated, if it affect other nodes. Based on the interestingness, the node is inserted into the tree.

The spreading tree represents the flow of information from root to leaves. OSNet is not possible to add dynamic seed nodes and imperceptible diffusion actions.

Ali et al. [5] proposed a signal processing framework, which can be used to summarize dynamic network by taking only representative networks. The framework creates key graph after finding the event intervals to summarize the intervals. The key graph that minimizes the cost function, which is measured as the mean squared error between the key graph and the other graphs, is selected. The summarization is done by convex optimization that uses quadratic programming. The summarization performance of the framework can be further improved by using different cost functions, algorithms for determining event interval and optimization methods.

IV. TEMPORAL NETWORK ANALYSIS MODEL IN DYNAMIC NETWORKS

Jiang et al. [6] proposed a temporal dynamic network analysis model and a multi-theoretical framework for examining different models in dynamic networks. It is an extension of multi-level, multi-theoretical framework for analyzing organizational networks. The framework proposes a set of temporal patterns which can be observed in online social networks and theoretical explanations on why these patterns are analyzed. The system proposes extended Exponential Random Graph model (ERGM) which is a statistical model that is useful for multi-dimensional networks. ERGMs are used to evaluate whether the observed networks exhibit theoretically hypothesized configurations.

The multi-theoretical framework includes configurations for testing temporal network patterns, which includes reciprocity standard, reciprocity temporal, co-occurrence standard, co-occurrence temporal, triangle standard, triangle temporal, k-star-standard and k-star-temporal. The system also proposes the hypothesis and selected relevant theories for these configurations. The proposed extended model and theoretical framework can be used for the deep understanding of the formation of online social networks. This framework is capable to extract some network dynamics that the proposed system cannot. This system can be used to any social networks that exhibit multi-dimensionality. If there is no knowledge about the network, then this model will be slow, which is a limitation. Furthermore, this work does not take into account of the nodal attributes, which particularly fit the ERGM for evaluating multi-dimensional social networks.

Jiang et al. [7] also proposed NATERGM (Nodal Attribute based Temporal Exponential Random Graph Model) to analyse the role of the attributes of node in tie formation. The model takes into account the order of tie formation, which is important to deeply understand the nature of a particular network. The order is analysed by taking the time details of the formation of ties. The model extends the ERGM model by adding temporal information into it. Thus it is capable of testing hypothesis regarding dynamic patterns. Different temporal variations of transitivity, reciprocity, K star, cyclicity are derived by considering the temporal information. The The parameter is estimated using Markov Chain Monte Carlo (MCMC) method. This model provides network prediction, by adjusting the parameters for network generation obtained from the differences in generated and actual network. The model is not able to consider multiple nodal attributes.

V. LATENT SPACE MODEL FOR DYNAMIC NETWORK ANALYSIS

Hoff *et al* [8] proposed a class of models where the probability actor's relationship depends on the positions of individuals in an unobserved "social space" are developed. The inference made in this paper is for the social space within maximum likelihood and Bayesian frameworks, and the system proposes Markov chain Monte Carlo procedures for making inference on latent positions and the observed covariate's effects. Using the latent position method, both directed and undirected relations can be analyzed. A conditional independence approach of modeling is adopted by assuming that the presence or absence of a tie between two individuals is independent of all other ties in the system.

Sarkar *et al* [9] proposed a method that uses latent space model to generalize dynamic model from static model. The model relates each entity with a point in Euclidean latent space. The links are observed between the entities, if they are close in latent space. An initial estimate of the position in latent space is obtained using Multidimensional Spacing (MDS). To this estimate Conjugate Gradient (CG) method is used for nonlinear optimization. This method could model the the evolving relationship between entities.

VI. PREVENTION OF INFERENCE ATTACKS IN SOCIAL NETWORKS

Heatherly *et al* [10] proposed methods that investigates the methods to launch inference attacks using the released data in social networking to predict private information about users that are not disclosed. The work devices three possible sanitization techniques that can be used in different situations. Attempts are made to use collective inference methods for discovering the attributes of the datasets that are sensitive. The work also discovers the problem domain that demean the performance of classification algorithm which determines private attributes. The social network is modeled as an undirected graph, where nodes represent the persons with a set of details and links denotes the friendship. A set of private details is further defined. Naïve Bayes is used as the learning algorithm to study the feasibility of possible inference attacks and the effectiveness of sanitization methods. The collective inference method is used to classify the social network which uses a combination of nodes and the links connecting them in the graph.

Four algorithms are proposed to predict the details of each user from the released data. They are "Details Only", "Links Only", "Average" and traditional Naïve Bayes classifier. The details are manipulated by adding details to nodes, modifying existing details and removing details from nodes. The links are also manipulated in the same way. The Naïve Bayes allows the easy scaling of the implementation to the large size of datasets. It also has the advantage of allowing simple selection methods to remove link and details while a network node class. The collective inference method controls the duration of that the algorithm runs. The work propose a best way to reduce the classifier accuracy and accuracy of local classifiers which results in maximum accuracy of the system. The work does not identify the key nodes that can greatly reduce the data leakage and also it ignores the temporal problem, which are the limitations.

VII. CONCLUSION

Dynamic Social Network Analysis is an emerging area that finds application in a wide variety of fields. Many algorithms have been proposed to address the challenges in this area. The methods for detecting communities/groups in dynamic networks and summarization of dynamic networks are explained. An effective framework for temporal network analysis and hypothesis testing is also mentioned. The paper also explains a method to prevent inference attacks on released data in social networks. The property of dynamicity makes the Dynamic Social Network Analysis, a different and tedious task than the conventional network analysis methods. This field is an emerging field that has a lot of chances to improve and explore.

VIII. ACKNOWLEDGMENT

I am extremely grateful to my PG Coordinator and guide Asst Prof. Salitha M. K , Assistant Professor of Computer Science and Engineering department Believers Church Caarmel Engineering College, Perunad, Ranni for her valuable guidance, timely suggestions and for providing all the vital facilities like providing the Internet facilities and important books, which were essential in the completion of this work. Finally, I would like to thank every individual who gave me even the slightest support to make my work a success.

REFERENCES

[1] C. D. Wang, J. H. Lai, and Philip. S, "NEIWalk: Community Discovery in Dynamic Content-based Networks," IEEE Trans. Knowl. Data Eng., vol. 24, no. 1, pp. 1734 - 1748, Jan. 2012.

[2] L. Tang, H. Liu, J. Zhang and Z. Nazeri, "Identifying Evolving Groups in Dynamic Multimode Networks," IEEE Trans. Knowl. Data Eng., vol. 6, no. 1, pp. 72 - 85, Jan. 2012.

[3] L. Tang, H. Liu and J. Zhang, "Community Evolution in Dynamic Multi-Mode Networks," ACM, KDD'08, pp. 24–27, August. 2008.

[4] Q. Qiang and L. Siyuan, Z. Feida, Z. Christian S. J., "Efficient Online Summarization of Large-Scale Dynamic Networks," IEEE Trans. Knowl. Data Eng., vol. 28, no. 12, pp. 3231 - 3245, Dec. 2016.

[5] Ali Y.M. and S. Aviyente, "Dynamic Network Summarization using Convex Optimization," IEEE Statistical Signal Processing Workshop., Jan. 2012.

[6] S. Jiang and H.Chen., "A Multi-Theoretical Framework for Hypothesis Testing of Temporal Network Patterns," 35th International Conference on Information Systems, Auckland, 2014.

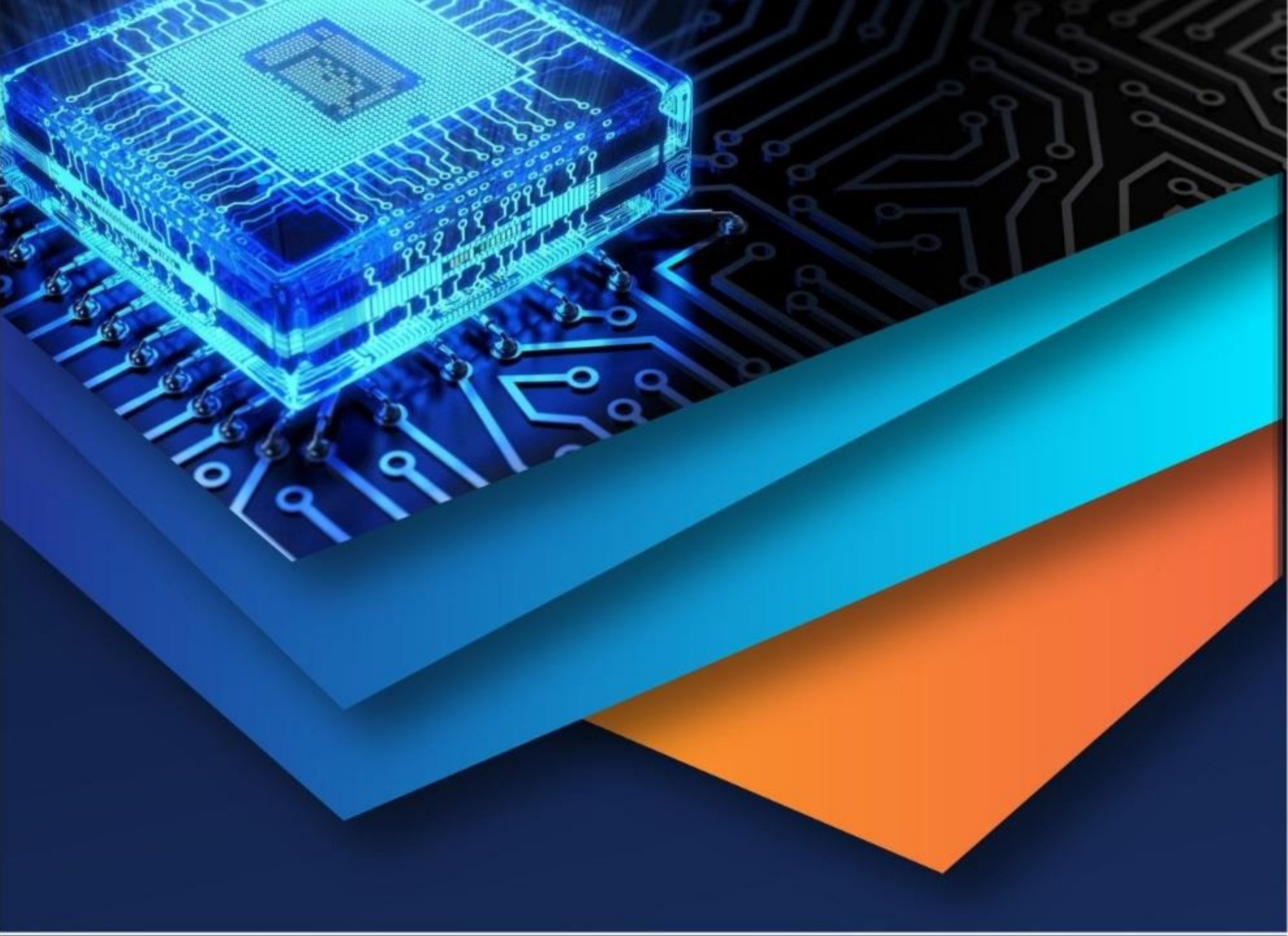
[7] S. Jiang and H. Chen., "NATERGM : A Model for Examining the Role of Nodal Attributes in Dynamic Social Media Networks," IEEE Trans. Knowl. Data Eng., vol. 28, no. 3, pp. 729 - 740, Mar. 2016.

[8] P. D. Hoff, A. E. Raftery, and M. S. Handcock., "Latent Space Approaches to Social Network Analysis", J. Amer. Statistical Assoc., vol. 97, no. 460, pp. 1090–1098, 2002.

[9] P. Sarkar, S. M. Siddiqi, and G. J. Gordon, "A Latent Space Approach to Dynamic Embedding of Co-occurrence Data", in Proc. 10th Int. Conf.Artif. Intell.Statist., 2007, pp. 420–427.

[10] R. Heatherly, M. Kantarcioglu, and B. Thuraisingham., "Preventing private Information Inference Attacks on Social Networks", IEEE Trans. Knowl. Data Eng., vol. 25, no. 8, pp. 1849 - 1862, Aug. 2013.

[11] https://en.wikipedia.org/wiki/Dynamic_network_analysis.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY