# DETECTING THE COMMUNITIES IN SOCIAL NETWORK USING ADAPTIVE DFFERENTIAL EVOLUTION

Jyotsna[1], Reena Rani[2]

[1] *student, computer science department, seth jai prakash mukundlal institute of technology,haryana, India*
[2] *Assistant professor, computer science department, seth jai prakash mukundlal institute of technology,haryana, India*

**ABSTRACT**

*Nowadays, several techniques are invented to detect communities in social network as numbers of people are interested in online social networks. Online communities are detected through cultural algorithm to solve the problem of optimization using knowledge repository but there are some limitations of this technique which is overcome by other method. This paper deals with the community detection in social network using adaptive differential evolution and compare the result with 4 well- known algorithms in this field and this method dramatically reduce the time of reaching to final answer.*

**Keyword**: *- communities, social network, adaptive differential evolution, and cultural algorithm.*

## 1. INTRODUCTION

Nowadays, networks like internet, we, social networks, biological networks, mobile and wireless networks become crucial part of human's life. These networks comprise of social agents and one can extract large amount of data using these networks. It is used in physics, computer science, biology and politics and sociology.

Network is nothing but a graph, G (V, E), where V is set of vertices and E is a set of edges that link the nodes. This graph is formed using adjacency matrix. Let A be an adjacency matrix then A (i, j) is 1 if there is a direct link between two nodes i and j and 0 if there is no edge.

Social network is the combination of connected communities. To understand the structure of the network, discovering the communities is vital. Therefore, various techniques and strategies have been conducted to find such communities efficiently. In short, a community is collection of nodes that has more internal links than external.

Community detection is a kind of NP- HARD problem and only evolutionary algorithm can solve this problem efficiently.

In this paper, we intend to detect community using adaptive differential evolution algorithm. The central idea of this algorithm comes from the fact that it detects the communities quickly, even faster than cultural algorithm, DE-CD and genetics. The central part of this algorithm is belief space which stores the information about sources and destinations. This belief space can be formed using different source of datasets.

The rest of the paper is organized as follows. In section 2, the main literature review is done in the same area. In section 3, merits and demerits of several algorithms are compared and contrast and finally the list of sources .

## 2. LITERATURE REVIEW

**Hao Lin (2014) [1]** Data mining tech like text categorization explore valuable information. Although energy consumed by text categorization inc. with the data. So, this energy is reduced by parallelization. Try to find classifies that perform best on effectiveness and efficiency. Classifies are obtained using existing library and implement classification algorithm. Improved versions of naïve Bayes can have competitive precision compared to SVM. Parallelization reduces energy cost by factor of 10 for RCVI dataset.

**Han Lu, Ananth Kalyanaraman et al (2014) [2]** Community detection is used in graph theoretic application. It reveals natural division that exists in real world network without knowing the size and cardinality of community. Community detection is rapidly detected by parallelization tech using '' louvain'' method. It is iterative heuristic for modularity optimization. This method is sequential thereby limiting its scalability. Heuristic is I/m using open MP multitasking and tested over internet, eitation, bio-logical. Compared to serial ''lovain'' method, parallel I/m produce community O/p with higher modularity. It speedups to 8 time using 32 threads. It exhibit weak scaling properties on up to 3 threads.

**Hassan Abbas Abdelbary, Abeer Mohamed ElKorany et al (2014) [3]** with increase in social network, users can identity other users with common interests, exchange their opinion. Discovering communities have become major challenges. Community detection approaches fall in 2 categories: - Consider user's network and utilizing user-generated content. A multiplayer community detection model applies gaussian restricted boltzman machine for making modeling user's post in network which yields to identify their topics of interest and finally construct communities. The effectiveness of this model is measured using KL divergence which measures similarity b/w users of same community.

**Fei Jiang, Jin Xu et al (2014) [4]** Existing approaches for community detection haven't considered about formation of communities. To detect overlapping and hierarchical communities, efficient and effective framework is used based on local influence. To demonstrate effectiveness, nodes with high local influence are selected to perform influence maximization experiments on real serial network. This framework is effective and efficient for both community detection and influence maximization. It illuminate two basic questions, whether local influence regarded as new property can effect formation of community or how to quantify node's local influence and use it to detect community. For this, nodes with high local influence are selected to perform influence maximization experiments on social network.

**Sanghamitra Bandyopadhyay, Garisha Chowdhary et al (2015) [5]** Overlap between communities are known to be denser most of the existing algorithm that detect overlapping communities assume that communities are dense and falsely identify overlap. Many algorithms do not scale with n/w size. An algorithm called fast overlapped community search (FOCS) is used to identify overlapped communities. FOCS is linear, gains in speed and outperform some popular overlapped community finding algorithm in terms of computational time not quality.

**Fang Li, Wentao Zhao et al (2015) [6]** A novel '' Modular-label-propagate-based'' algorithm is used to discover high quality community. This algorithm starts with propagation of network vertexes and then binds node with compact structures. Greedy method is applied which work according to sequence of neighbor mode size, if module value increases, sequence will be renewed. This algorithm improves quality effect and stability of found communities to be close to linear complexity.

**Michal Prilepok, Milos Kudelka (2015) [7]** Network community analysis algorithm is applied to detect spam by finding communities and then grouping the e-mails. New nearest community classifier is used and the result is obtained very close to Baysian spam filter. Accuracy is 93.78% and detects 80.77% of spam emails and 98.01% of non spam emails.

**Murat Can Gainz, Melike Tutka et al (2015) [8]** Supervised meaning classifier (SMC) based on Helmholtz principle from gestalt theory is used for text classification. New SMC outperform traditional classifier of multinomial naïve bayes (MNB) and support vector machine (SVM).

**Yongliang Wu, Li He et al (2015) [9]** an algorithm is developed to detect overlapping community in complex network. Community detection using this algorithm is more accurate and has lower time complexity as compared to traditional methods.

**E. Raju, M.A Hameed et al (2015) [10]** novel method called ''non normalized spectral clustering '' incorporated with bisecting K-means is used to identify communities. The experiment is performed on three real world datasets Zachary karate club, American college football and bottle nose dolphin network. The experiments show that the method is effective. This algorithm performs better than benchmark algorithm.

**Nidheesh Melethadathil, Priya Chellaiah et al [11]** neuro informatics natural language processing relies on clustering and classification. Role of text classification and document clustering algorithm on datasets, where features were obtained by mapping to manually verified. NLP data classification involves eight techniques and unsupervised learning was performed with six clustering algorithm.

**Han Hu, Yonggang Wen et al [12]** social video replication and users request dispatching in cloud CDN architecture used to reduce operational cost of system. Community classification method cluster social users with social relationship. Conduct large scale measurement on real OSN system to study diversities of social video propagation and request dispatching strategy and formulate is as constrained optimization problem.

**P. Anupriya, S. Karpagavallin [13]** topic modeling is powerful technique for analysis large document collection. Top models have wide range of application like tag recommendation, text categorization, keyword extraction and similarity search in broad field of text mining. The documents models are built using LDA (latent dirichlet allocation) with collapsed variation bayes and Gibbs sampling. Gibbs sampling outperforms CVBO sampling.

**Zongqing Lu [14]** most of the existing community detection algorithms focus on binary network. It proposes inter centrality and inter-centrality, to characterize nodes in communities, based on which they propose data forwarding algorithm for DTN and worm containment strategy for OSN.

**Pooya Moradian Zadeh and Ziad Kobti 2015 [15]** during the last decade community detection is hotly debated issue. They developed a knowledge based evolutionary algorithm that solved the problem of complex network using multi population cultural algorithm. They first extract the knowledge from network and then find its optimal solution. This algorithm rapidly finds out the true communities more precisely.

**Leonardo N. Ferreira and Liang Zhao 2015 [16]** they developed time series clustering technique based on community detection. Here a network of nodes and vertex is constructed where vertex represent the time series. Time series distance functions calculate the similarities between vertex data. Lastly, time series clusters are calculated using community detection algorithm. This algorithm is effective and outperformed many old algorithms.

**Liangxun Shuo and Bianfang Chai 2015 [17]** they tried to summarize evolution law of algorithm. They also tried to improve the existing community detection algorithm using statistical inference.

**Wiliam J. Salter 2015 [18]** he suggested that community detection is linked with real world problem based on real data sets than focusing on technical advances. He concluded suggestions of research areas and approaches that could further progress toward this goal.

**Pooya Moradian Zadeh and Ziad Kobti 2015 [19]** they proposed an algorithm called cultural algorithm to solve complex problems to get optimal result. This algorithm is compared with other four known algorithms and results outperform in terms of time and find communities in fewer generation.

**Arif Mahmood and Michael Small [20]** most exciting community detection algorithms are graph theoretic and lack the ability to detect accurate community boundaries if ratio b/w intra to inter community is low. Algorithm based on modularity maximization to resolve common smaller than the specific size. So, author present different community detection algorithm based on fact that each n/w common spans diff subspaces in geodesic space. To make community detection more robust, sparse linear coding with norm constraints are

used. Sparse spectral clustering algorithm is used to label each node. For this, nodes with high local influence are selected to perform influence maximization experiments on social n/w.

## 3. COMPARISON AND CONTRAST

The comparisons of different research are as follows:

| Sr. No. | AUTHORS | TITLE | PROS | CONS |
|---|---|---|---|---|
| 1. | Hao Lin | Research on Energy Efficient Text Classification | This method is precise and reduced total energy cost | Only parallelization is used to save energy. |
| 2. | Hassan Abbas Abdelbary | Utilizing deep learning for content based community detection. | Their model gives best result by reducing dimension of dataset for single layer community detection. | Increased computational cost. |
| 3. | Fei Jiang, Shuyuan Jin, Yanlei and Jin XU | A uniform framework for community detection via influence maximization in social network. | Test on both real network and artificial benchmarks have resulted in accurate and effective algorithm for overlapping and hierarchical algorithms. | |
| 4. | Sanghamitra Bandyopdhyay | FOCS: fast overlapped community search | It explores communities fast and reflects real world communities.<br><br>Maximum overlap is set. | The maximum number of communities that can be examined is equal to the number of nodes in network. |
| 5. | Fang Li, Wentao Zhao, Zhifeng Sun and Bin Dong | A kind of community detecting algorithm based on modularized label propagation. | Gets communities of higher quality. It is more consistent in reality. | |
| 6. | Michal Prilepok | Spam detection on nearest community classifier | Powerful spam filter as compared to Bayesian spam filter.<br><br>80.72% spam email and 98.01% non spam are detected. | Accuracy is lower in comparison with Bayesian spam filter. |
| 7. | Murat Can Gainz, Melike Tutkan and Selim Akyokus | A novel classifier based on meaning for text classification | It uses name measure to calculate meaning in terms of context.<br><br>It outperforms other text classification algorithm. It is faster. | No work on negative meaning values.<br><br>Semantic relation is not defined properly. |

| | | | | |
|---|---|---|---|---|
| 8. | Yougliang Wu, Li He | An algorithm for overlapping community detection in complex network | Accurate and lower time complexity.<br><br>Communities that are detected have clearer meaning. | |
| 9. | E. Raju | Detecting communities in social networks using un normalized spectral incorporated with bisecting K-means. | Effective and give better result on all datasets. | Cannot detect automatically the number of clusters in a network. |
| 10. | Nidheesh Melethadathil, Priya Chellaiah, Bipin Nair and Shyam Diwakar | Classification and clustering for Neuroinformatics: assessing the efficacy on reverse-mapped neuroNLP data using standard ML techniques. | Preprocessing the data set by reverse mapping improve the quality of mined set and their inter-relations. | Reverse mapping is used |
| 11. | P. Anupriya and S. Karpagavalli | LDA based topic modeling of Journal Abstracts. | Extracted topics using LDA capture meaningful data. | |
| 12. | Zongqing Lu | Algorithms and applications for community detection in weighted networks. | Better result and outperforms other algorithm. | |
| 13. | Arif Mahmood | Subspace based network community detection using sparse linear coding. | To make process robust. Sparse linear coding is used. It is consistent. | |
| 14. | Hao Lu and Ananth Kalyanaraman | Parallel heuristics for scalable community detection | High modularity and exhibit weak scaling properties. | Do not work on large scale inputs. |

## 4. CONCLUSIONS

It is further concluded that there are various ways of detecting communities in social network using different methods but if communities are detected using adaptive differential evolution then it outperform in terms of time.

## 6. REFERENCES

[1]. Abbas Abdelbary Hassan, Abeer Mohamed ElKorany and Reem Bahgat "*utilizing deep learning for .content based community detection*" 2014.
[2]. Bandyopdhyay sanghmitra, Garisha Chowdhary and Debakar Sengupta "*FOCS: fast overlapped community search*" vol, 27 No. 11, November 2015.
[3]. Hu han"*joint content replication and request routing for social video distribution over cloud CDN: a community clustering method*" 2015.
[4]. Gainz M. C, Melike Tutkan and Selim Akyokus "*a novel classifier based on meaning for text classification*" 2015.
[5.] Jiang Fei, Shuyuan Jin, Yanlei Wu and Jin Xu "*a uniform framework for community detection via influence maximization in social network"* 2014.
[6.] Li Fang, Wentao Zhao, Zhifeng Sun and Bin Dong *"a kind of community detecting algorithm based on modularized label propagation"* 2015.
[7.] Lin Hao "*research on energy efficient text classification*" 2014.

[8.] Lu Hao, Ananth Kalyanaraman, Mahantesh Halappanavar and Sutanay Choudhury *"parallel heuristics for scalable community detection"* 2014 DOI 10.1109, IPDPSW.2014.155
[9.] Lu Zongqing *"algorithms and applications for community detection in weighted networks"* 2015.
[10.] Melethadathil nidheesh, priya chellaiah, bipin nair and shyam diwakar *"classification and clustering for neuroinformatics: assessing the efficacy on reverse- mapped neuroNLP data using standard ML techniques"* 2015.
[11.] Mahmood Arif *"subspace based network community detection using sparse linear coding"* 2016.
[12.] Michal Prilepok and Milos Kudelka *"spam detection based on nearest community classifier*" DOI 10.1109/INCoS.2015.75
[13.] P. anupriya and S. Karpagavalli "*LDA based topic modeling of journal abstracts*" 2015.
[14.] Wu Yongliang, Li he, Guanghui Yan, Fanglin Guo, Weitao Zheng and Abdul Basit Khan "*an algorithm for overlapping community detection in complex network"* 2015.
[15.] Zadeh P. M. and Ziad Kobti "*community detection in social network using cultural algorithm*" 2015.