International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Community Structure Analysis using Fast Louvain Method in Real World Networks

Laxmi Chaudhary, Buddha Singh, Neeru Meena

Abstract: Recently, in complex networks detection of Community structure has gained so much attention. It adds a lot of value to social, biological and communication networks. The community structure is a convoluted framework thus analyzing it helps in deep visualization and a better understanding of complex networks. Moreover, it also helps in finding hidden patterns, predicting link in various types of networks, recommending a product to name a few. In this context, this paper proposes an agglomerative greedy method, referred to as Fast Louvain Method (FLM), based on Jaccard cosine shared metric (JCSM) to deal with the issues of community structure detection. Specifically, Jaccard cosine shared metric (JCSM) is developed to find the similarity between the nodes in a network. We have utilized modularity quality function for assessing community quality considering the local changes in this network. We test the method performance in different real-world network datasets i.e. collaboration networks, communication networks, online social networks, as well as another miscellaneous networks. The results also determined the computation time for unveiling the communities. This proposed method gave an improved output of modularity, community goodness, along with computation time for detecting communities' number as well as community structure. Extensive experimental analysis showed that the method outperforms the existing methods.

Keywords: community detection, community structure, complex networks, modularity, real world networks, social networks

I. INTRODUCTION

The traffic of social networks is rapidly increasing every year. There is an enormous growth in the online intercommunication of users. In recent times various online social networking platforms like Twitter, Facebook and many more have provided user intercommunication as a result to raise interactions, it is a challenge to keep the track of user interaction data. These real-world online social networks have intriguing patterns and characteristics, which can be scrutinized for various objectives [1].

Online Social Networks have an essential characteristic called community structure. Discovering community structure has been in hotspot for past few years. Communities have dense links among themselves and sparse links having left network [2]. Identifying prominent nodes for targets in viral marketing [3], recommending a product for online retail sites [4], finding hidden pattern in a network and predicting a link are the most significant applications of community detection. Therefore, finding and analyzing communities can greatly contribute in understanding the large and convoluted structures of social networks. Hence different techniques have been introduced to detect communities in the networks with each having certain drawbacks, such as graph partitioning methods [2], [5], [6] (spectral and Kernighan-Lin methods), Hierarchical clustering approaches [7] (Agglomerative and Divisive) and modularity optimization techniques [7], [8], [9], [10], [11] (Greedy, Extremal optimization, Simulated annealing).

Recent development in community detection is hinge to optimize a quality function proposed by Girvan along with Newman [4] called as modularity modelled depends upon the random null hypothesis assuming that community structures do not exist in random networks. Modularity examines merits of a partition within the network. Finding the community with the highest modularity is a tedious task that was already prove to be NP-hard [12]. Several modularity-based techniques like the Blondel algorithm [13] have been developed to find a local optimal partition. This technique detects communities based on the modularity gain, and the size of network is restricted because of limited storage capacity instead of limited computational time. PyLouvain method [14] for community detection is particularly depends upon the concept of shared edges.

In this particular research paper, we have conceptualized FLM (Fast Louvain Method) to explore community structure of complex networks. The significance of given proposed technique is mention below:

- We defined a greedy technique which operates in an agglomerative hierarchical manner to find communities for small to large datasets using modularity metric score [15] for learning communities.
- We introduced a new similarity measure, Jaccard Cosine shared metric (JCSM) to measure the relationship among the nodes.
- Prior knowledge of an existing network structure is not required.
- The proposed method detects better communities as compared to the existing approaches as the obtained value of modularity metric is high.
- In the network the computation time taken for discovering the communities reduced significantly.

Manuscript published on November 30, 2019.
* Correspondence Author
Laxmi Chaudhary*, SC & SS, Jawaharlal Nehru University, New Delhi, India. Email: laxmichaudari.iet@gmail.com
Buddha Singh, SC & SS, Jawaharlal Nehru University, New Delhi, India. Email: b.singh.jnu@gmail.com
Neeru Meena, SC & SS, Jawaharlal Nehru University, New Delhi, India. Email: neerumeena15@gmail.com

We tested the performance of the algorithm using different real-life network datasets viz. communication network, collaboration network, and social network along with another miscellaneous network datasets. The time required for simulation was computed by the newly introduced technique to reveal the network communities. Inspired by Louvain [13], and PyLouvain [14] method, we used modularity metric score to examine the essence of the communities. The evaluation of a modularity metric score depicts that the newly introduced method shows convincing reformation in comparison to the prior techniques.

The paper has been designed in Section 2. FLM (Fast Louvain method) is defined in section 3. Whereas, Section 4 presents verification of obtaining outcomes on several different datasets of real-world network. The conclusion of this research paper along with deriving some prudent ways to move further is provided in 5th Section.

II. RELATED WORK

Lately, a concept of plethora was developed in a network to uncover the community structure. Community structure detection technique related to modularity optimization have opened up a new scope in complex network systems. The greedy method is used to construe the attributes of networks in community detection analysis. Researchers from disparate areas such as physics, social science, statistics, data mining, and computer science are engaged in community structure detection for quite some time now and have put forward various algorithms in complex networks to reveal communities.

Newman and Girvan [15] presented a divisive approach in a complex network to derive community structure. This particular technique depends upon the edge betweenness centrality measure for detecting communities. Edge Betweenness for a link is a number of shortest paths (geodesic paths), actually joins node-pair passing through the link. The main objective of this approach is to find link with high edge betweenness and then construct the communities by removing these edges of the original network. As we know, edges that connects two groups have high Edge Betweenness. If more than two edges have the same highest Betweenness then it randomly selects any edge to delete, or simultaneously deletes all of edges. Horizontal cut at any level of the dendrograms created at the end of this technique depicted the possible many groups present in this network depending on the cut's position. The main drawback was that this technique did not provide any details about the communities of the network divided by it.

Radicchi et. al. [16] suggested a method same as to that of Girvan and Newman [15] algorithm, the only difference being that it was based on edge clustering coefficient measure instead of Edge Betweenness. In Girvan and Newman [15], we delete the link between the high Edge Betweenness value in every iteration, but in the algorithm developed by Radicchi et. al. in whole iteration, the link having smallest edge clustering coefficient is removed. This coefficient defined as the number of triplets consisting the edge divided by the total number of possible triplets that may be made. The drawback of the method is its high computing time. Newman [17] developed an agglomerative method in the network for discovering the communities based on modularity score. The modularity is illustrated as the difference among these connections within these modules along with expected number of connections randomly connected among the 2 nodes in a graph. In this approach, a group of nodes is successively merged to form bigger groups so that the modularity value rises after the integration taking quadratic computational time to unveil communities in networks.

Clauset et. al. [18] also developed a technique based on the modularity value to measure the community structures. The algorithm uses max load data structure and an ordinary vector array. At the end of this method, a dendrogram, a tree is constructed representing the order in which the nodes are combined and cut at any level of this dendrogram revealing the communities in the network. The cut might be chosen by looking for maximum modularity score. Computational time is the main limitation of this method.

Schuetz et. al. [19] introduced a multistep greedy algorithm (MSGA) which is also an agglomerative hierarchical method. In MSGA, more than a pair of groups are merged at every iteration step. This multistep simple refinement process known as vertex mover which is utilized to reassign nodes to neighbouring groups to enhnace the final modularity score. Again, this technique also requires more time for computation.

Danon et. al. [20], developed an approach to unveil groups in a network. This method is a small modification of the Newman [15] method that does not consider the heterogeneity of community size. It favored the formation of a bigger group at the expense of smaller groups. The Danon et. al. algorithm considered communities of various sizes and treated them equally. In this method, the distance measure is replaced by a change in modularity score. Though this approach outperforms Newman method but the computation time is still high for this technique.

Blondel et. al [13], presented an algorithm known as the Louvain technique to find communities. This particular algorithm is an efficient pathway to hinge on the agglomerative hierarchical technique, in which the group of nodes are repeatedly connected to form large modules in such a way to increase the modularity gain [15] after merging. The partitions of the large networks with high modularity score are obtained in a short time revealing the hierarchical structure of the communities by constructing a dendrogram. This method is designed to detect a local optimal community. The algorithm dwells in two phases which are reprised iteratively until the communities in a network structure are revealed. When the algorithm completes two phases, it is considered that the algorithm is complete pass. The computational time is again the drawback of this method as well.

V.A. Traag [21], introduced an approach to speed up the Blondel algorithm. This method considered moving nodes to a random neighbour community, rather than the optimal neighbour community. The computational time of this method is slightly greater than the Louvain method which runs in linear time concerning the number of links.

Julien et. al. [14], proposed a community detection method, known as PyLouvain method. In the PyLouvain method, this method is depending upon the Louvain algorithm [13] presenting a shared edge concept to disclose communities in a network.

In this context, FLM (Fast Louvain method) is presented to reveal the network's community structure by introducing a new notion, Jaccard Cosine shared metric (JCSM) to find the similarity between nodes. We also compared the simulation time taken by the proposed FLM method and other existing methods.

III. PROPOSED FAST LOUVAIN METHOD (FLM)

This section presents the proposed "Fast Louvain Method" (FLM) to unveil communities in real world networks. This method reveals a complete community structure in an agglomerative hierarchical manner. FLM, is a greedy technique to depict the communities in the network. The proposed method derives partitions of large networks by utilizing the JCSM (Jaccard Cosine shared metric) along with modularity with less computation time. The FLM is divided into three steps that repeat iteratively.

I) Similarity Computation
II) Modularity Optimization
II) Community Structure Aggregation

A. First Step-Similarity Computation

In the first step of FLM, each node is considered to be an individual community in a network. Consequently, in this network the number of nodes will be the number of communities. Further, considering that each node is represented with *i* along with *c* is the number of neighboring nodes. This proposed technique, *i* node may change with its *c* number of neighbors implying that node *i* has *c* communities for its movement but, node *i* chooses one of the best neighbors for its shift. For finding the finest neighbor node *i* moves with whole neighbors, for every shift of node *i* we compute the Jaccard Cosine Shared Metric (JCSM). Therefore, we obtain *c* JCSM number values corresponding to whole neighboring network node *i*. *i* node shifts to its neighbor community depends upon the JCSM's highest value. This particular step is employed iteratively for all single node till no further enhancement can be obtained.

Jaccard Cosine Shared Metric

We used two similarity measure i.e. Cosine measure [22] and Jaccard measure [23]. To identify the correspondence among the nodes Cosine Similarity [22] is utilized as it is also a similarity measure. This similarity measure helps in effectively determining the similarity between nodes in large networks. It considers the sparse nodes present in the network with low complexity as compared to other measures. The cosine similarity Sim_{jg} [22], of a moving node *j* to community *g* is estimated by

$$Sim_{jg} = \left(\frac{j.g}{||j||\,||g||}\right) \quad (1)$$

Jaccard Similarity [23] is utilized to obtain the similarity among nodes. It is defined as the ratio between neighborhoods' intersection and union. This measure efficiently checks which nodes are shared and which are distinct. We define a new JCSM which combines the features of Cosine measure, and Jaccard measure. We also consider the average shared links *r* to provide more accurate closeness value among nodes. JCSM, J_{ig} of *j* moving node to *g* community is estimated by

$$J_{jg} = \left(\frac{|\Gamma(j)\cap\Gamma(g)|}{|\Gamma(j)\cup\Gamma(g)|}\right) + Sim_{jg} + r \quad (2)$$

where $\Gamma(i)$ and $\Gamma(c)$ are the neighbourhoods of nodes *j* along with *g*. Average shared links is represented by r.

$$r = \left(2 * s - \frac{d_j * d_g}{l}\right) \quad (3)$$

where shared links between the nodes *j* and *g* represents by S. Degree of nodes *j* and *g* represents by d_j along with dg. Number of links among these nodes in the graph represents by *l*.

B. Second Step-Modularity Optimization

After applying the first step of the FLM method, we obtained the communities depends upon on JCSM, which gave the similarity score among the nodes. In second step of FLM approach, we calculate the modularity gain of the partition. There are various objective functions other than modularity like Potts models, infomap to name a few. The modularity score can be calculated by the modularity [15] objective function. The community structure having the maximum modularity gain [15] is a finest community structure. *Q* (Modularity score) [15], is calculated by

$$Q = \frac{E.W}{2l} - \left(\frac{d_j d_g}{(2l)^2}\right) \quad (4)$$

where *E.W* is the edge weight of linking node *j* and node *g*.

C. Third Step-Community Structure Aggregation

In the third step, we construct a new network of communities in such a way that the groups achieved during the second step are expected as a node for forming the network of communities. In third step, the new formed network, weights of edge amongst the nodes is compared as the sum of the edge weights within the community. The edges achieved between similar groups led to self-loops.

After the third step finished, all the steps of the developed technique are applied again to the resulting network and repeated until the community structure having high modularity score is not obtained. After obtaining community structure, this proposed approach computes time taken by each network to reveal the communities.

The algorithm of the FLM method is described below.

Algorithm (FLM)

Step 1. Assume each node as a community.

- *Calculate the JCSM (J_{ig}) of each node.*
- *Connect the node to the adjacent node community on the maximum value of J_{ig}*
- *Iterate above steps until all the nodes of a network join the community.*

Step 2. Find the modularity score (Q).

Step 3. Compute edge weights (E.W) of all the nodes in each community.

- *Edges among same community become self-loops.*
- *Construct the community network.*

Step 4. Go to Step 1 Repeat until the maximum modularity score is not obtained.

Step 5. Calculate the time taken for obtaining communities.

Step 6. Stop the algorithm.

IV. EXPERIMENT AND RESULT ANALYSIS

Here we evaluate the newly introduced FLM using 14 real world networks. The FLM method, is tested after comparing with Louvain technique [13] as well as PyLouvain technique [14]. The proposed method is programmed in Python 3.4.4, a computer having an Intel Core i7, 12 GB RAM along with 3.20 GHz.

The list of parameters used in proposed FLM is presented in Table I. The goodness of the community structure is calculated by utilizing the modularity function [15]. Table II provides the brief detail of a real-world networks.

Table I. List of Parameters

Parameters	Descriptions
n	Number of nodes
l	Number of links
Q	Modularity score
C	Number of communities
E. W	Edge Weight
$d_j, d_j, d_g d_g$	Degree of nodes j and g
r	Average shared links

Table II. Real World Networks

S. no.	Networks	Nodes	Links	Description
1.	Karate [24]	34	77	Zachary's karate club network
2.	Soc-firm-hi-tech [25]	33	147	Social Network
3.	Dolphin [26]	62	159	Dolphin Network
4.	Lesmis [27]	77	254	Les Miserables Network
5.	Football [28]	115	613	Football Network
6.	Email [29]	1133	10903	Email network
7.	Email enron [30]	36692	367662	Email enron network
8.	CA-AstoPh [31]	18772	396160	Astro physics network
9.	CA-HepPh [31]	12008	237010	High energy physics network
10.	CA-Gr [31]	5242	28980	General relativity and quantum cosmology network
11.	CA-HepTh [31]	9877	51971	High energy physics theory network
12.	CA-CondMat [31]	23133	186936	Condense matter network
13.	Arxiv [32]	9377	48214	Arxiv Network
14.	Citations [32]	27770	352807	Paper citations network

A. Simulation Results

The performance of the newly introduced FLM method is assessed on 14 famous real-world datasets. On these real-world networks the obtained results demonstrated that newly introduced FLM method performs well in relation to another existing approaches in number of communities along with computation time and modularity. Table III represents the experimental results obtained from testing various datasets. It shows that the modularity score metric value of the FLM method has significantly improved as compared to Louvain and PyLouvain techniques. This signifies that the number of communities obtained using proposed method are superior. It also shows the computation time taken to obtain the communities by utilizing Louvain along with PyLouvain method, and FLM method.

Table- III: demonstrates the number of communities along with Computation time, modularity score, taken to obtain communities using Louvain along with PyLouvain method and the FLM technique. The modularity obtained by FLM has shown improvement as compared to other approaches.

S.no	Network	Louvain Method			PyLouvain Method			FLM Algorithm		
	Name	c	Q	$T(s)$	c	Q	$TT_{(s)}$	c	Q	$T(s)$
1.	**Karate [24]**	4	0.430	0.042	4	0.430	0.058	4	0.440	0.016
2.	**Soc-firm-hi-tech [25]**	4	0.316	0.087	4	0.317	0.085	4	0.338	0.067
3.	**Dolphin [26]**	5	0.538	0.067	5	0.519	0.067	5	0.547	0.050
4.	**Lesmis [27]**	6	0.550	0.078	6	0.550	0.074	6	0.570	0.059
5.	**Football [28]**	9	0.604	0.162	10	0.604	0.161	8	0.617	0.150

6.	**Email [29]**	10	0.566	0.800	12	0.564	0.855	11	0.574	0.751
7.	**Email enron [30]**	1279	0.602	98.700	1284	0.605	99.900	1262	0.620	91.102
8.	**CA-AstoPh [31]**	393	0.862	103.120	393	0.862	103.990	393	0.885	99.10
9.	**CA-HepPh [31]**	490	0.772	43.121	479	0.767	44.437	475	0.787	39.110
10.	**CA-Gr [31]**	324	0.627	2.311	321	0.628	2.344	324	0.643	1.59
11.	**CA-HepTh [31]**	321	0.656	6.533	325	0.659	6.535	316	0.676	4.520
12.	**CA-CondMat [31]**	618	0.731	50.124	618	0.731	50.261	622	0.744	47.551
13.	**Arxiv [32]**	59	0.813	5.144	62	0.814	5.147	59	0.823	4.140
14.	**Citations [32]**	170	0.643	70.501	167	0.643	68.761	167	0.660	62.521

Fig. 1 shows the improvement in the modularity score of all the datasets for the proposed FLM in comparison to the existing techniques. Consequently, for the proposed method the quality of community structure is also good in comparison with Louvain along with PyLouvain technique. Fig. 2 demonstrates the number of communities detected using proposed FLM method and other methods on various datasets. Fig. 3 shows the modularity score analysis to the number of nodes using FLM method in comparison with other methods. It represents that the modularity score value of the FLM approach is higher as the number of nodes increases. Fig. 4 shows the modularity score variation to the number of by utilizing the FLM technique in comparison with PyLouvain and Louvain method.

Fig. 5 represents the computation time taken to reveal the structure of community of the various datasets, using proposed FLM and other methods. The evaluation time for determining the community structure also improved these networks. Therefore, the performance of the developed FLM method is better than existing approaches.

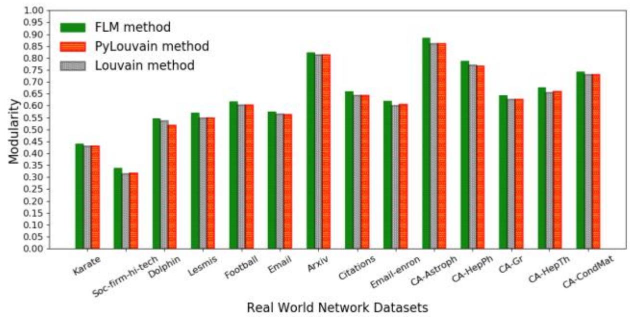

Fig. 1. Comparison of the Modularity score over various datasets by utilizing Louvain technique along with PyLouvain technique and FLM method.

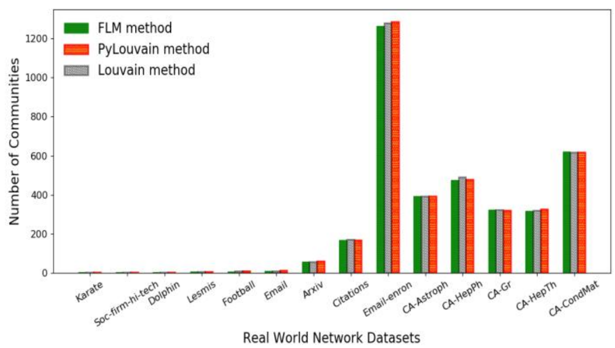

Fig. 2. Number of Communities over Datasets by utilizing Louvain technique along with PyLouvain technique and FLM method.

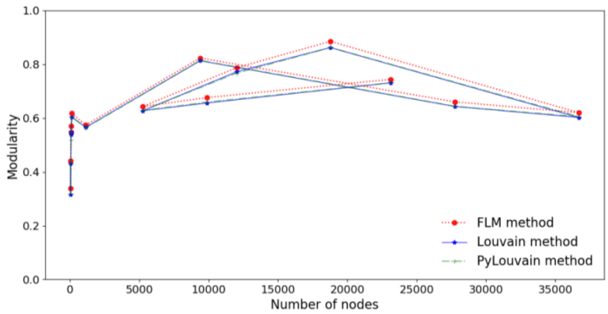

Fig. 3. Comparison of Modularity score metric versus number of nodes by utilizing Louvain technique along with PyLouvain technique and FLM method.

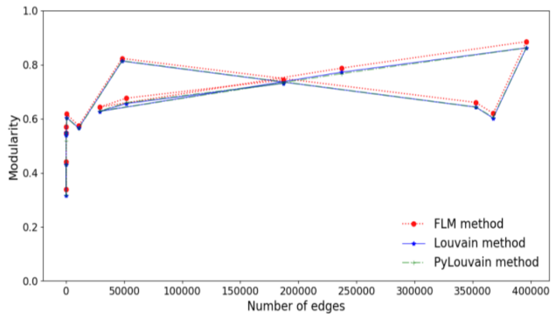

Fig. 4. Comparison of Modularity score metric versus number of edges by utilizing Louvain technique along with PyLouvain technique and FLM method.

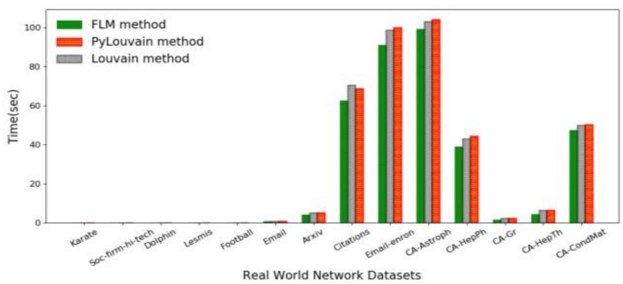

Fig. 5. Evaluation of Time over Datasets by utilizing Louvain technique along with PyLouvain technique and FLM method.

V. CONCLUSION

This particular research paper introduced a greedy technique in complex datasets, known as Fast Louvain method (FLM), for revealing the community structure.

This particular method works in an agglomerative manner for exploring communities in various networks. The proposed FLM method extracts the similarity measure and the value of Jaccard Cosine shared metric (JCSM) is used to uncover the hidden pattern of the network. Additionally, the modularity rating is enhanced by allowing just regional modifications of communities in the system. The time taken for finding the communities was also computed for the performance evaluation. The performance of FLM method was evaluated for 14 real world datasets like; communications networks, collaboration networks, online social network, and another miscellaneous networks, which showed that the FLM approach is robust and more significant as compared to another existing technique. The obtained results illustrated the outstanding performance of the developed FLM method in comparison with the past approaches in modularity score as well as time taken to extract the structure of community. The modularity score is compared with the number of edges along with nodes. The modularity score is also analyzed with these datasets. Also, the number of communities have been analyzed in comparison with the number of nodes along with these datasets. From the evaluation of all the obtained results it may be concluded that the proposed FLM technique betters the existing approaches. For future work, another parameter of the nodes such as density will be considered, and employed to elevate the similarity among these nodes for exploiting the communities' quality detected in this network.

REFERENCES

1. Bedi P, Sharma C. Community detection in social networks. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2016; 6:115-135.
2. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. Physical Review E 2004; 69:026113.
3. Combefis S. Viral marketing and Community Detection Algorithms. Universite Catholique de Louvain Faculte des sciences appliquees Departement dingenierie informatique 2007.
4. Sahebi S, Cohen WW. Community-based recommendations: A solution to the cold start problem. In Proceedings of the Workshop on Recommender Systems and the Social Web, RSWEB (RecSys'11). ACM, Chicago 1997.
5. Newman MEJ. Finding community structure in networks using the eigenvectors of matrices. Physical review E 2006; 74:036104.
6. Newman MEJ. Detecting community structure in networks. The European Physical Journal B-Condensed Matter and Complex Systems 2004; 38:321-330.
7. Fortunato S. Community detection in graphs. Physics reports 2010; 486:75- 174.
8. Newman MEJ. Modularity and community structure in networks. Proceedings of the national academy of sciences 2006; 103:8577-8582.
9. Guimera R, Sales-Pardo M, Amaral LAN. Modularity from fluctuations in random graphs and complex networks. Physical Review E 2004; 70:025101.
10. Duch J, Arenas A. Community detection in complex networks using extremal optimization. Physical review E 2005; 72:027104.
11. Chaudhary L., Singh B. (2019) Community Detection Using an Enhanced Louvain Method in Complex Networks. In: Fahrnberger G., Gopinathan S., Parida L. (eds) Distributed Computing and Internet Technology. ICDCIT 2019. Lecture Notes in Computer Science, vol 11319. Springer, Cham
12. Brandes U, Delling D, Gaertler M, G¨orke R, Hoefer M, Nikoloski Z, Wagner D. Maximizing modularity is hard. arXiv preprint physics/0608255 2006.
13. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment 2008; 10:10008.
14. Odent J, Guillain MS. Automatic detection of community structure in networks. 2012.
15. Girvan M, Newman MEJ. Community structure in social and biological networks. Proceedings of the national academy of sciences 2002; 99:7821-7826.
16. Radicchi F, Castellano C, Cecconi F, Loreto L, Parisi D. Defining and identifying communities in networks. Proceedings of the National Academy of Sciences of the United States of America 2004; 101:2658-2663.
17. Newman MEJ. Fast algorithm for detecting community structure in networks. Physical review E 2004; 69:066133.
18. Clauset, Aaron, Mark EJ Newman, and Cristopher Moore. "Finding community structure in very large networks." Physical review E 70.6, p. 066111, (2004).
19. Schuetz P, Caflisch A. Efficient modularity optimization by multistep greedy algorithm and vertex mover refinement. Physical Review E 2008; 77:046112.
20. Danon L, Díaz-Guilera A, Arenas A. The effect of size heterogeneity on community identification in complex networks. Journal of Statistical Mechanics: Theory and Experiment 2006; 11:11010.
21. Traag VA. Faster unfolding of communities: Speeding up the Louvain algorithm. Physical Review E 2015; 92:032801.
22. Singhal A. Modern Information Retrieval: A Brief Overview. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 2001; 24:35–43.
23. Jaccard P, Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. Bulletin de la Société Vaudoise des Sciences Naturelles 1901; 37:241-272.
24. Zachary WW. An information flow model for conflict and fission in small groups. karate club Network. Journal of Anthropological Research 1977; 33:452-473.
25. Rossi R, Ahmed N. The Network Data Repository with Interactive Graph Analytics and Visualization. AAAI. 2015;15.
26. Lusseau D, Schneider K, Boisseau OJ, Haase P, Slooten E, Dawson SM. Behavioral Ecology and Socio-biology 2003;54:396-405.
27. Knuth DE. The Stanford GraphBase: A Platform for Combinatorial Computing. Addison-Wesley 1993.
28. Girvan M, Newman MEJ. Network of American football games. Proc. Natl. Acad. Sci. USA 2002; 99:7821-7826.
29. Guimera R, Danon L, Diaz-Guilera A, Giralt F, Arenas A. Physical Review E 2003; 68:065103.
30. Leskovec J, Lang KJ, Dasgupta A, Mahoney MW. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. Internet Mathematics 2009; 6:29-123.
31. Leskovec J, Kleinberg J, Faloutsos C. Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data 2007;1.
32. Leskovec J, Kleinberg J, Faloutsos C. Graphs over time: densification laws, shrinking diameters and possible explanations. Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining. ACM 2005.

AUTHORS PROFILE

Laxmi Chaudhary received her M. Tech degree in Computer Science from School of Computer and Systems Sciences at Jawaharlal Nehru University, New Delhi, India in 2016. Currently, she is a Ph.D. research scholar at School of Computer and Systems Sciences, Jawaharlal Nehru University. Her research interest includes Community detection, Social Networks and Artificial Intelligence.

Buddha Singh received his B.Tech degree from Madhav Institute of Technology and Science, Gwalior, India and Ph.D degree from Jawaharlal Nehru University, New Delhi, India. In 2014, he joined Jawaharlal Nehru University as an Assistant Professor. His current research interest includes Mobile Ad-hoc Network, Wireless Sensor Network, Cognitive Radio, Big Data Analytics, Complex Networks, Mobile Computing.

Neeru Meena received her M. Tech degree in Computer Science from School of Computer and Systems Sciences at Jawaharlal Nehru University, New Delhi, India in 2016. Currently, she is a Ph.D. candidate at School of Computer and Systems Sciences, Jawaharlal Nehru University. Her research interest includes Wireless Sensor Networks.