Proceedings on Big Data Analytics
& Innovation (Peer-Reviewed),
Volume 1, 2016, pp.110-123

An Enhanced Conductance-Based Approach for Community Detection in Weighted Mobile Phone Networks

ELIZABETH N. ONWUKA, BALA A. SALIHU & PASCHAL S. IORNENGE
Federal University of Technology, Minna, Nigeria

ABSTRACT Community Detection has gained a lot of attention in recent years due to its applications in studying human behaviour in various spheres of life and most especially in the analysis of criminal networks. In this era of Big Data Analytics, community detection has been made easier by the availability of huge sources of data such as the Call Detail Records (CDR) of the telephone networks. Recently, focus in community detection is gradually drifting from unweighted networks to weighted networks, where the strength of the link between each pair of connected nodes is considered rather than just the existence of a link. However, existing algorithms for community detection have focused only on direct links between pairs of nodes in a network. In this work, an Enhanced Conductance-based Algorithm (ECBA) was develop to detect communities in a network. This was done by synthesizing the direct and indirect relationship strengths between all pairs of nodes on a weighted undirected graph. The algorithm was tested with CDR data using belonging degree and conductance as the decision metrics for community partitioning. Comparison with the original conductance-based algorithm shows significant improvement in quality of detection for communities of large sizes in terms of average shortest path distances, density, and how closely knit the connections are. Test results further show that using indirect relationships between pairs of nodes significantly reveals more information about community membership in large networks

Keywords: community detection; social network graphs; binary networks; weighted networks; conductance; belonging degree.

1. Introduction

The massive improvement in technology over the years, especially the technical and commercial success of the mobile phones and other handheld devices, has made the study of human behaviour or interaction patterns via this medium increasingly useful. Users of mobile phones, either for voice/SMS communication or for online social networks, leave digital traces that can be used to understand their behaviour and connections over time. Even criminal activities and criminal networks can be more easily understood and detected by analysing such data. Among the prominent ways this has been done is community detection.

A network is a group of nodes (or vertices) connected through edges or links. A community in a network is a group of nodes having more internal connections with

each other than external connections with the rest of the network (Fortunato, 2010). They are also called Clusters, Cliques or Cohesive groups (Borgatti, 2009) (Palla et al., 2005). Detecting communities of users on a network has gained significant growth due to applications such as warm containment in Online Social Networks (OSN), data forwarding in delay tolerant networks, routing strategies for MA-NETS, coding, automatic allocation of small LANs (Lu et al., 2013), detecting terrorist or criminal groups, Link prediction, information diffusion, and in biological or medical systems (Ahuja et al., 2016). This is commonly done on social network graphs which are made of nodes (users on the network), and edges which represent links between the users. Social Network graphs may be directed or undirected, weighted or unweighted. In a directed graph, the direction of communication between two nodes is considered, while Undirected graphs are made up of unordered pairs of vertices, i.e., direction of communication is not used in carrying out analyses. An undirected graph is unweighted (or binary) if a single edge connects each pair of vertices. An unweighted graph (or an unweighted version of a graph) is used for analyses when the goal is to simply know which nodes have communication links to each other. In this case there is no interest in the extent of communication in those links. However, for weighted graphs, there can be multiple edges connecting a pair of vertices, highlighting the extent of communication and hence, the relative strengths of the links (Lu et al., 2013).

In this paper, we approach the problem of community detected on weighted and undirected networks. Our main contribution is the development of an Enhanced Conductance-based Algorithm (ECBA) which not only uses direct relationship strengths but also indirect relationship strengths to improve the quality of community detection.

2. Related Works

The goal of community detection is to partition a network into dense regions of the graph. Each region represents a group of nodes that are closely related, and hence are in the same community. Most of the earlier algorithms for community detection were based on binary networks. Very prominent among them is one proposed in (Girvan and Newman, 2002) which focused on the boundaries of communities rather than their core. It is said to be the first algorithm in modern age community detection (Fortunato and Lancichinetti, 2009). In their approach edges are removed from the network based on Betweeness centrality values. The edges with the highest Betweeness centrality are removed, Betweeness is calculated again for the edges affected by this removal, and the process is repeated until no edges remain. However, the run time of the algorithm as the number of nodes increase makes it unsuitable for large graphs. Cfinder was developed to uncover the structure of complex networks by analysing the statistical features of overlapping networks (Palla et al., 2005). A community (a k-clique community) was defined as a union of all k-cliques (complete subgraphs of size k) that can be reached from each other through a series of adjacent k-cliques (where adjacency means sharing $k-1$

nodes) $k-1$. It was based on the fact that members can be reached through well connected subsets of nodes. This approach allowed overlapping in community membership. The community detection was done by setting a threshold weight for the links and ignoring links that were below this threshold weight making the network essentially a binary network. The RAK algorithm which is based on label propagation was also proposed in (Raghavan et al., 2007). In this approach, each node is first initialized to a unique label which represents the community it belongs to, and these labels then propagate through the network. A node would determine its community based on the labels of its neighbours. Each node joins a community which has the most of its neighbours as members and the labels of the nodes are updated at each iteration. As the propagation continues, dense connected groups of nodes finally settle for a unique label, and in the end, all nodes with the same labels are placed in the same community. This continues until each node in the network has the label to which the maximum number of its neighbours belong to. It is however possible for the iteration to end with two disconnected groups of nodes having the same label. It will require a breadth-first search on the subnetwork of each individual group to separate the disjointed communities thus increasing the computation time and complexity of the technique.

All of the methods briefly discussed earlier focused on binary networks. In such methods, attributes of nodes are emphasized instead of the edge content which represent the actual link between the nodes. Even though more challenging, edges provide a richer characterization of community behaviour (Qi et al., 2012). Most networks are weighted, so community detection is more reliable when the actual extent of interaction between nodes is considered (Ovelgönne et al., 2010). A notable algorithm for detection of communities in weighted networks is the COPRA algorithm (Gregory, 2010). This algorithm is based on the label propagation algorithm (RAK) discussed earlier. Label propagation is done just like in the RAK algorithm only that, in this case, a node can be a member of more than one community because of the use of community identifiers. A node is allowed to keep more than one community identifier in each label without retaining all of them. This algorithm can be used on weighted networks. However, it has the same convergence problem that the RAK algorithm had. In (Tiantian Zhang and Bin Wu, 2012) a method for finding communities of users by first identifying core nodes and finding cliques around those core nodes was proposed. It was argued that having global knowledge of the graph required by most algorithms is unrealistic for very large graphs. The *Strength* algorithm proposed in (Chen et al., 2010) used this strategy. It consists of finding an initial partial community (the node with the highest node strength). The community is expanded by adding tight nodes to the partial community until detection is complete for that particular community based on a set threshold for belonging degree of the neighbours of that community. The algorithm however, degrades in its performance when the overlapping increases. In (Lu et al., 2013), a conductance-based algorithm was developed. The algorithm is just like the *Strength* algorithm only that a new objective function, Conductance, is used in addition to the belonging degree, and here the initial community is a community of two nodes in the network with the highest edge weight between the two of them.

This algorithm had a dynamic threshold and could perform well on large networks, unlike the *Strength* algorithm. However, like all the previous algorithms discussed, indirect links between nodes were not considered.

Sometimes, friends (and also criminals like fraudsters) live in close proximity to each other. This reduces the amount of communication data available to study their relationships since most of their communication happen offline (Blackburn et al., 2014). Considering indirect connections can help to reveal more information about such relationships. According to Granovetter (1973), "The degree of overlap of two individual's friendship networks varies directly with the strength of their tie to one another." Thus, nodes with stronger ties to each other are more likely to have stronger indirect links or friends-of-friends. In an attempt to determine the distance in a communication network beyond which two nodes are no longer likely to be aware of each other's activities in (Friedkin, 1983), it was observed that two persons who were more than two steps away from each other in a network were unlikely to be aware of each other's work. Work by Christakis and Fowler, (2009) also led to a theory that social influence does not end with two people who are directly connected to each other but continues up to two or three hop relationships, though with diminishing returns. Work carried out by Blackburn et al., (2014) further verified this.

3. Community Detention with Synthesised Relationship Strengths

Here, we present a method for detecting communities using a synthesised relationship strength (direct and indirect) between pairs of directly connected nodes in a network.

A. Synthesised Relationship Strengths

In (XLin et al., 2014), a simple expression for calculating synthesized relationship strengths between pairs of nodes in a network was derived. The synthesized relationship strength is the weight of the link between any two nodes, it is derived from the combination of the weights of the direct the indirect paths between the two nodes. The synthesized relationship strength $RS(v_i, v_j)$ between nodes v_i and vj is written as

$$RS(v_i, v_j) = \propto RS_d(v_i, v_j) + \beta RS_{id}(v_i, v_j)$$

$$RS(v_i, v_j) = \propto RS_d(v_i, v_j) + \beta RS_{id}(v_i, v_j) \quad (1)$$

Where $\propto$ and β a $\propto$ re we β ighting coefficients for the direct and the indirect paths respectively. Selecting the experimental values for the attenuation coefficient and weight coefficient as used in (XLin et al., 2014), The synthesized relationship strength $RS(v_i, v_j)$ between nodes v_i and v_j

$$RS(v_i, v_j) = 0.6w_{i,j} + 0.4\frac{\sum_{i=1}^{n}\left(\prod_{j=1}^{d_j} w_j\right)}{n}$$

$$RS(v_i, v_j) = 0.6w_{i,j} + 0.4\frac{\sum_{i=1}^{n}\left(\prod_{j=1}^{d_j} w_j\right)}{n}$$

(2)

where $w_{i,j}$ repres $w_{i,j}$ ents the direct weight between the two nodes, and d is the length of their relationship strength along a given path (number of hops in between)

For c two-hop indirect paths with intermediary node v_k , where v_k v_k is a n v_k eighbour of both v_i and v_j , v_i the s v_j um of weights, P_1 across P_1 all such indirect paths was calculated as:

$$P_1 = \sum_{1}^{c}[(w_{i,k}] \times w_{k,j})$$

$$P_1 = \sum_{1}^{c}[(w_{i,k}] \times w_{k,j})$$

(3)

For m three-hop indirect paths with intermediary nodes v_k and v_l , v_k where v_l v_k is a n v_k eighbour of v_i , v_l is a v_i n v_l eighbour of v_k , and v_j i v_k s in a v_j neighbour of v_l ; the s v_l um of weights, P_2 across P_2 all such indirect paths was calculated as:

$$P_2 = \sum_{1}^{m}[(w_{i,k}] \times w_{k,l} \times w_{l,j})$$

$$P_2 = \sum_{1}^{m}[(w_{i,k}] \times w_{k,l} \times w_{l,j})$$

(4)

Therefore,

$$RS_{id}(v_i, v_j) = 0.4\frac{(P_1 + P_2)}{(c+m)}$$

$$RS_{id}(v_i, v_j) = 0.4\frac{(P_1 + P_2)}{(c+m)}$$

(5)

Hence,

$$RS(v_i, v_j) = 0.6w_{i,j} + 0.4\frac{(P_1 + P_2)}{(c + m)}$$

$$RS(v_i, v_j) = 0.6w_{i,j} + 0.4\frac{(P_1 + P_2)}{(c + m)} \quad (6)$$

This was then used in place of $w_{i,j}$ in the $w_{i,j}$ conductance-based algorithm.

B. Metrics Used to Detect Communities

The following metrics were used as objective functions in our algorithm for detecting communities.

1) Conductance: It measures the fraction of total edge volume that point outside the cluster. That is, it measures how well knit a graph is. The lower the conductance value, the more connected the nodes are. This can be mathematically expressed as:

$$\phi(C) = \frac{cut\left(C, C/G\right)}{w_c}$$

$$\phi(C) = \frac{cut\left(C, C/G\right)}{w_c} \quad (7)$$

where $cut\left(C, C/G\right)$ repres $cut\left(C, C/G\right)$ ents the number of cut edges in the community (which means all edges leaving the community), and w_c is the w_c total weight of edges in the community.

2) Belonging Degree: Assuming C is a community in a network; for a node $u \in V$; k_u , N_u , a $u \in V$ re k_u n N_u ode degrees and neighbour sets respectively. And let w_{uv} be the w_{uv} weight of the link between nodes u and v (where *v* is already in the community). k_u can th k_u en be written as:

$$k_u = \sum_{u \in N_u} w_{uv}$$

$$k_u = \sum_{u \in N_u} w_{uv} \quad (8)$$

For the community C, and node u, the belonging degree $B(u, C)$ betwee $B(u, C)$ n node u and community C is defined as

$$B(u, C) = \frac{\sum_{u \in C} w_{uv}}{k_u}$$

$$B(u, C) = \frac{\sum_{u \in C} w_{uv}}{k_u}$$

(9)

C. The ECB Algorithm

The algorithm is made up of two stages: selecting the initial temporary community and expansion. It is basically the Conductance-based Algorithm (CBA) in (Lu et al., 2013), with synthesized relationship strengths used in place of direct weights between all links in the entire process. The algorithm is as follows

(a) Input Graph G

(b) Ca G lculate synthesized relationship strength between every pair of nodes in the network

(c) If edge set is not empty, select two nodes v_i and v_j w v_i ith t v_j he highest synthesized relationship strength

(d) Calculate the Conductance $\Phi(C)$ of the $\Phi(C)$ community C forme C d by v_i and v_j

(v_i e) Fi v_j nd all the neighbours (N) of C

(f) Pi C ck the neighbour with the highest belonging degree B(w, C) to C

(g) Ad C d to C and fo C rm a new temporary community C'

(h) Ca C' lculate conductance $\Phi(C')$ of C'

(i $\Phi(C')$) If C' $\Phi(C') < \Phi(C)$, the $\Phi(C') < \Phi(C)$ n $C' = C$, go to $C' = C$ (e)

(j) If $\Phi(C') > \Phi(C)$, then $\Phi(C') > \Phi(C)$ community C is det C ec-ted

(k) Remove edge v_i - v_j , go t v_i o v_j (c)

(l) End

4. Performance Evaluation of the ECBA

The algorithm was tested on two datasets—i.e. the nodobo dataset and ground truth data from Zachary Karate Club.

1) Test with Nodobo Dataset

The Nodobo dataset (McDiarmid et al., 2013) is publicly available. This data was retrieved from mobile phones of 27 High School students over a period of 5 months using a software. It consists of 13035 call records, 83542 SMS records and 5.2 million proximity records. The part of the dataset used for this research is the call records. Based on the scope of this work, only the source and the target phone numbers and call durations were needed. These were extracted and duplicate edges (source-destination pairs) were merged. Since this work focuses on undirected graphs node pairs were considered as duplicates if they existed as a source-destination pair, however they were permutated, multiple times. Weights were calculated using call durations for each edge.

$$\text{Weight} = \frac{(\text{Call duration with respect to a given edge})}{(\text{Maximum call duration in the dataset})}$$

This pre-processing resulted in 575 nodes with 642 edges.
The result of the communities detected are shown as compared to that of the original Conductance-based Algorithm as shown in Table 1.

Table 1: Detected communities

Number of members	Detected by ECBA	Detected by CBA
> 5	23 communities	23 communities
>10	16 communities	18 communities
>20	11 communities	12 communities
>30	8 communities	9 communities
>40	6 communities	7 communities
>50	2 communities	3 communities
>60	1 community	1 community
>70	1 community	0 communities

From the results, it is clear our algorithm generally detected more small-sized communities than the existing algorithm. It showed more details of splits among members. Also the largest community detected by our algorithm had 105 nodes. The same community (with the same initial node pair), was detected with only 68 nodes using the CBA algorithm. Thus, with our algorithm, it was possible to identify

members of that community which were not detected by the CBA. Other metrics used for evaluation include:

a) Conductance

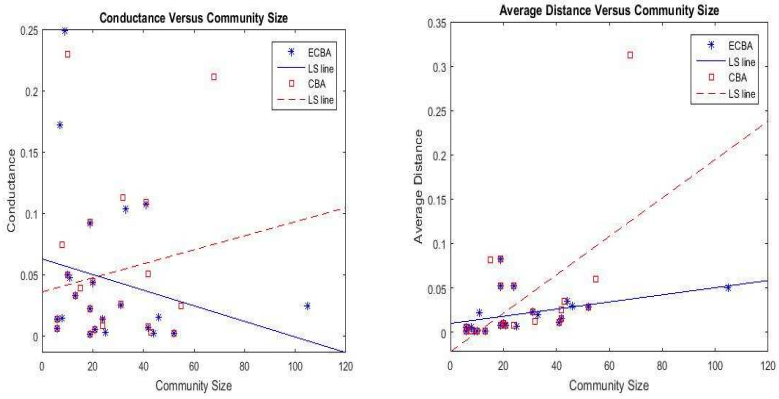
The performance of the Enhanced Conductance-based Algorithm (ECBA) was compared with the Conductance-based Algorithm (CBA) by plotting the graphs of their conductance versus community size. The result shows a scatter plot with Least Squares lines in Fig 1a. As seen from the least-square lines, the ECBA had lower conductance values as community sizes increased from about 20 members and above. By definition, the smaller the conductance, the tighter the connection (that is, the stronger the relationships) between members in the community. This means that the ECBA formed tighter clusters for communities with larger sizes (greater than 20) while the CBA formed tighter clusters for smaller community sizes. The last two scatter points to the right show that the community detected by the ECBA with 105 nodes is much tighter (conductance = 0.02456) than that of the same community which was detected as having only 68 nodes by the CBA (conductance = 0.2112). Hence, the hidden nodes ignored by the CBA were very important members of the community.

b) Average Distances

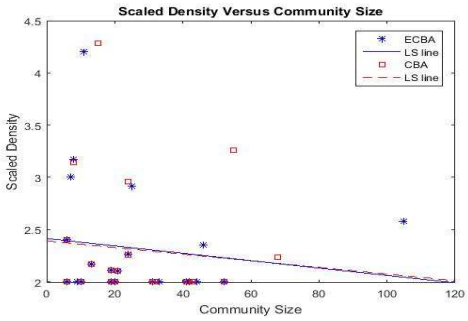
The distance between two nodes in a network is the length of the shortest path between them. For this network, the distances used were inverse of the weights as discussed earlier. The average distance is the sum of the shortest path distances between all pairs of nodes in the community divided by the community size. Fig 1b shows a plot of the average distances for each of the communities formed using the ECBA and the CBA against community sizes. The ECBA finds communities with lower average distances as the sizes of the communities are increasing, while the CBA shows lower average distances for smaller communities (<20). This can be interpreted as the ECBA clustering nodes with a stronger connection with one another (smaller shortest path lengths) than the CBA for larger communities.

c) Average Distances

The values of the scaled densities for each of the communities detected by both algorithms were plotted against community sizes in Fig 1c. The ECBA has a higher scaled density up to point 40 on the community size axis and CBA has a higher scaled density from point 80 upwards. However, the scatter points show only ECBA having a community at all up to 80 nodes in size. Hence we can best compare the two algorithms with earlier values than 80. This shows that the ECBA communities have averagely slightly higher density of clusters than the CBA.



(a) (b)



(c)

Fig 1: Graphs comparing performance of the ECBA and CBA

2) Test with Zachary Dataset

Our algorithm was also tested with the Zachary Karate club dataset. This data is from an already known community structure of 34 members of a Karate club observed for three years - from 1970 to 1972. After a conflict between the club's president and a part-time instructor over lesson fees, members of the club were split into two main groups (Zachary, 1977). In this work, a weighted version of the karate network was used to test the Enhanced Conductance-based Algorithm. The results of the detection were compared with the real life communities observed by Zachary. The original communities formed by Zachary are shown in Table 2.

Table 2: Community Structure as observed by Zachary

Community	Members
1	1, 2, 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 17, 18, 20, 22
2	9, 10, 15, 16, 19, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34

The detection done with our algorithm shows four communities instead of the two observed by Zachary as seen in Table 3. However, it was observed that members of community 2 and community 4, except for node 9, 29 and 32, were both subsets of the community 1 in Zachary's original detection. Also, community 1 and community 3 in Table 2, except for node 2 and 3, are subsets of community 2 in Zachary's detection.

Table 3: Community Structure Observed by the ECBA

Community	Members
1	25, 26, 29, 32
2	1, 2, 3, 4, 8, 12, 13, 14, 18, 20, 22, 29
3	2, 3, 9, 10, 14, 15, 16, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34
4	6, 7, 17, 1, 9, 32, 5, 11

When the pairs of subsets are merged we have the communities as shown in Table 4. It can be seen from comparing Table 2 and Table 4 that every member of the community 1 in the Zachary dataset is also in the first community we have after merging the community pairs as shown in Table 4. Also, every single member of community 2 in the Zachary dataset is also placed in the second community in Table 4. The extra nodes in each of these communities, that is: 2, 3, 9, 14, 20 and 29, are overlapping nodes.

Table 4: Merged Communities

Community	Members
2 & 4	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 20, 22, 29, 32
1 & 3	2, 3, 9, 10, 14, 15, 16, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34

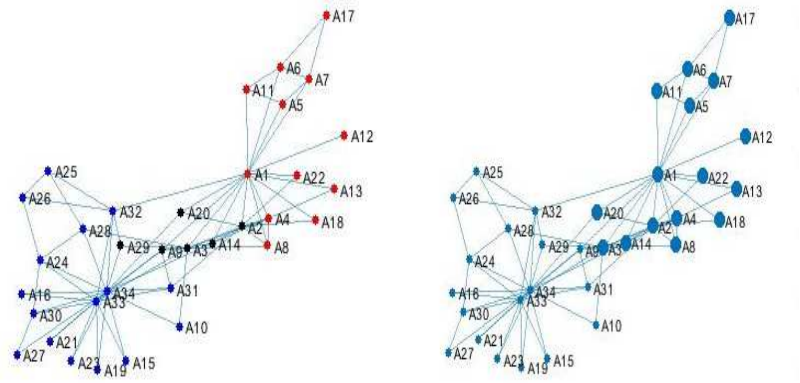


Fig 2: (a) Communities as observed Zachary (b) Communities as detected by ECBA

The result of this test shows that the proposed algorithm did not only detect communities but also sub-communities within these communities, showing more detailed clustering of nodes.

5. Discussion

The results from the detection and evaluation discussed in the previous section shows that more community members not seen by the CBA can be detected by the ECBA. The results point to the fact that using indirect relationship strength compensates for some of the closely connected nodes that have very little online communication and could hence be mistaken for weak ties. Apart from showing the true nature of such links, communities formed showed a great improvement in compactness. This also suggests that a higher amount of mutual information is shared across the links between members of communities formed via synthesising both direct and indirect paths. The test with Zachary karate club dataset also shows that the introduction of indirect relationships across the links gave rise to overlapping. Though every member was detected in its correct community, overlaps became very visible and sub-communities could be formed to show greater detail of how members related.

6. Conclusions

In this work, an Enhanced Conductance-based Algorithm (ECBA) for community detection in weighted networks with undirected graphs was presented. It was tested on a mobile phone dataset and a dataset from a social club with already known community structure. Results show that The Enhanced Conductance-based Algorithm outperforms the existing Conductance-based Algorithm in detecting communities of larger sizes (up to about 20 nodes or more). This work has also revealed that detecting communities with both direct and indirect relationship strengths

gives more details of node relationships than what is obtained by using only direct relationship strengths.

Correspondence
Bala A. Salihu
Department of Telecommunications Engineering
Federal University of Technology, Minna, Nigeria

References

Ahuja, M., Singh, J. & Neha, 2016. Practical Applications of Community Detection. *International Journal of Advanced Research in Computer Science and Software Engineering,* 6(4), pp. 412-415.

Blackburn, X.Z.J., Kourtellis, N., Skvoretz, J., Iamnitchi, A., 2014. The power of indirect ties in friend-to-friend storage systems, in: 14-Th IEEE International Conference on Peer-to-Peer Computing. IEEE, pp. 1–5.

Borgatti, S.P., 2009. 2-Mode concepts in social network analysis. Encyclopedia of complexity and system science 6.

Chen, D., Shang, M., Lv, Z., Fu, Y., 2010. Detecting overlapping communities of weighted networks via a local algorithm. Physica A: Statistical Mechanics and its Applications 389, 4177–4187. doi:10.1016/j.physa.2010.05.046
Christakis, N.A., Fowler, J.H., 2009. Connected: the surprising power of our social networks and how they shape our lives, 1st ed. ed. Little, Brown and Co, New York.

Friedkin, N. E., 1983. Horizons of observability and limits of informal control in organizations. *Social Forces,* 62(6), pp. 54-77.

Fortunato, S., 2010. Community detection in graphs. Physics Reports 486, 75–174. doi:10.1016/j.physrep.2009.11.002

Fortunato, S., Lancichinetti, A., 2009. Community detection algorithms: a comparative analysis: invited presentation, extended abstract, in: Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), p. 27.

Girvan, M., Newman, M.E., 2002. Community structure in social and biological networks. Proceedings of the national academy of sciences 99, 7821–7826.

Granovetter, M. S., 1973. The Strength of Weak Ties. *American Journal of Sociology,* 78(6), pp. 1360-1380.
Gregory, S., 2010. Finding overlapping communities in networks by label propagation. New Journal of Physics 12, 103018. doi:10.1088/1367-2630/12/10/103018.

Lu, Z., Wen, Y., Cao, G., 2013. Community detection in weighted networks: Algorithms and applications, in: Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on. IEEE, pp. 179–184.

McDiarmid, A., Bell, S., Irvine, J., Banford, J., 2013. Nodobo: Detailed mobile phone usage dataset. Unpublished paper, accessed at http://nodobo. com/papers/ietel. pdf on 9–21.

Ovelgönne, M., Geyer-Schulz, A., Stein, M., 2010. Randomized greedy modularity optimization for group detection in huge social networks, in: Proceedings of the Fourth SNA-KDD Workshop, KDD 2010, July. pp. 1–9.

Palla, G., Derényi, I., Farkas, I., Vicsek, T., 2005. Uncovering the overlapping community structure of complex networks in nature and society. Nature 435, 814–818. doi:10.1038/nature03607.

Qi, G.-J., Aggarwal, C.C., Huang, T., 2012. Community detection with edge content in social media networks, in: 2012 IEEE 28th International Conference on Data Engineering. IEEE, pp. 534–545.

Raghavan, U.N., Albert, R., Kumara, S., 2007. Near linear time algorithm to detect community structures in large-scale networks. Physical Review E 76. doi:10.1103/PhysRevE.76.036106.

Tiantian Zhang, Bin Wu, 2012. A Method for Local Community Detection by Finding Core Nodes. IEEE, pp. 1171–1176. doi:10.1109/ASONAM.2012.202.

XLin, X., Shang, T., Liu, J., 2014. An Estimation Method for Relationship Strength in Weighted Social Network Graphs. Journal of Computer and Communications 02, 82–89. doi:10.4236/jcc.2014.24012.

Zachary, W.W., 1977. An information flow model for conflict and fission in small groups. Journal of anthropological research 452–473.