Sink Group Betweenness Centrality

Evangelia Fragkou
University of Thessaly
Volos, Greece
efragkou@e-ce.uth.gr

Dimitrios Katsaros
University of Thessaly
Volos, Greece
dkatsar@e-ce.uth.gr

Yannis Manolopoulos
Open University of Cyprus
Nicosia, Cyprus
yannis.manolopoulos@ouc.ac.cy

ABSTRACT

This article introduces the concept of Sink Group Node Betweenness centrality to identify those nodes in a network that can “monitor" the geodesic paths leading towards a set of subsets of nodes; it generalizes both the traditional node betweenness centrality and the sink betweenness centrality. We also provide extensions of the basic concept for node-weighted networks, and also describe the dual notion of Sink Group Edge Betweenness centrality. We exemplify the merits of these concepts and describe some areas where they can be applied.

CCS CONCEPTS

• **General and reference** → **Metrics**; **Evaluation**; • **Human-centered computing** → **Collaborative and social computing design and evaluation methods**; **Social network analysis**; *Social engineering (social sciences).*

KEYWORDS

Sink Group Betweenness Centrality, Sink Group Edge Betweenness Centrality, Betweenness Centrality, Network Science

ACM Reference Format:

Evangelia Fragkou, Dimitrios Katsaros, and Yannis Manolopoulos. 2021. Sink Group Betweenness Centrality. In *25th International Database Engineering & Applications Symposium (IDEAS 2021), July 14–16, 2021, Montreal, QC, Canada.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3472163.3472182

1 INTRODUCTION

The dramatic growth of online social networks during the past fifteen years, the evolution of Internet of Things and of the emerging Internet of Battlefield Things, the extensive study and recording of large human social networks is offering an unprecedented amount of data concerning (mainly) binary relationships among ‘actors’. The analysis of such graph-based data becomes a challenge not only because of their sheer volume, but also of their complexity that presents particularities depending on the type of application that needs to mine these data and support decision making. So, the field termed *network science* has spawned research in a wide range of topics, for instance: a) in network growth, developing models such as the preferential attachment [4], b) in new centrality measures for the identification of the most important actors in a social network developing measures such as the bridging betweenness centrality [28], c) in epidemics/diffusion processes [18], d) in finding community structure [21, 22], i.e., finding network compartments which comprise by sets of nodes with a high density of internal links, whereas links between compartments have comparatively lower density, e) in developing new types of networks different from static and single-layer networks, such temporal [39] and multilayer networks [9], and of course analogous concepts for these types of networks such as centralities [26], communities [27], epidemic processes [5], and so on.

Despite the rich research and the really large number of concepts developed during the past twenty years and the diverse areas where is has been applied e.g., ad hoc networking [30], the field of network is continuously flourishing; the particular needs of graph-data analytics create the need for diverse concepts. For instance, let us examine a very popular and well-understood concept in the analysis of complex networks which is the notion of centrality [3, 34, 41], being either graph-theoretic [23], or spectral [33] or control theoretic [29]. Betweenness centrality [23] in particular has been very successful and used for the design of effective attacks on network integrity, and also for discovering good “mediators" (nodes able to monitor communication among any pair of nodes in a network), but it is not effective in identifying influential spreaders; for that particular problem k-shell decomposition [31] proved a much better alternative. However subsequent research [11] proved that a node’s spreading capabilities in the context of rumor spreading do not depend on its k-shell index, whereas other concepts such as the PCI index [6] can perform better. So, our feeling is that the field of network science, even some very traditional concepts such as betweenness centrality could give birth to very useful and practical variants of them.

Let us describe a commonly addressed problem in network science concerning malware spreading minimization problem. In particular, we are given a set of subsets of computer nodes that we need to protect against a spreading malware which has already infected some nodes in the network, but we only have a limited number of vaccines (i.e., a “budget") to use. If we had to select some healthy (susceptible) nodes to vaccinate, then which would these nodes be? Our decision would of course depend on the infection spreading model, but usually routing in computer networks implements a shortest-path algorithm. So can the traditional shortest-path node betweenness centrality [23] help us to identify such nodes?

Additionally, we will describe a similar problem that a modern army could possible face due to the rapid deployment of Internet of Battlefield Things [43]. The army needs to destroy by jamming the communications towards a set of subset of nodes of the enemy, but

IDEAS 2021, July 14–16, 2021, Montreal, QC, Canada

ACM ISBN 978-1-4503-8991-4/21/07.
https://doi.org/10.1145/3472163.3472182

it has available only a limited number of jammer or time to deploy them. The question is now which links should the army select to attack? So can the traditional shortest-path edge betweenness centrality [15] help us to identify such links?

The answer to previous questions is negative, because node/edge betweenness centrality calculates the importance of a node/edge lying on many shortest paths connecting any pair of network nodes, whereas in our problems we are interested in paths leading towards specific sets of subsets of network nodes. Starting from this observation, in this article we introduce the generic concept of *sink-group betweenness centrality* which can be used as a building block for designing algorithms to address the aforementioned problems.

The aim of the present article is to introduce the new centrality concept for various types of complex networks and also to present some potential uses of it for solving some network science problems. In this context, the present article makes the following contributions:

- it introduces a new centrality measure, namely the sink group node betweenness centrality;
- it extends the basic definition for networks weighted on the nodes;
- it extends the basic definition for the case of edge betweenness;
- it presents basic algorithmic ideas for calculating the aforementioned notions.

The rest of the article is structured as follows: Section 2 presents the articles which are closely related to the present work; Section 3 defines the sink group betweenness centrality concept, and in Section 4 we provide a detailed example to exemplify the strengths of the new concept. Then, in Sections 5 and 6 we present the definitions of sink-group betweenness centrality for node-weighted networks and edge sink group betweenness centrality, respectively. Finally, Section 7 discusses a list of applications and Section 8 concludes the article.

2 RELATED WORK

Betweenness Centrality.

The initial concept of (shortest path) betweenness centrality [23] gave birth to concepts such as the proximal betweenness, bounded-distance betweenness and distance-scaled betweenness [13], bridging centrality [28] to help discover bridging nodes, routing betweenness centrality [16] to account for the paths followed by routed packets in a networks, percolation centrality [38] to help measure the importance of nodes in terms of aiding the percolation through the network. There are so many offsprings of the initial concept, that even a detailed survey would found practically impossible to record each one published!

The realization that the computation of betweenness centrality requires global topology knowledge and network-wide manipulation, which is computationally very expensive, spawned research into distributed algorithms for its computation [10], and inspired variants aiming at facilitating the distributed computation of notions similar to the original betweenness centrality, such as load centrality [36].

Betweenness centrality and its variations have found many applications not only in classical fields in network mining, but also in delay-tolerant networks [37], ad hoc networks [30], in distributed systems e.g., for optimal service placement [42], etc.

Approximate Betweenness Centrality.

Many applications working over modern networks require the calculation of betweenness centrality in nearly a real-time fashion or have to deal with a huge number of nodes and edges. So, the decision of trading off the accuracy of betweenness centrality computation with speed arises as a natural option. Some early works considered approximating the exact values of betweenness centrality [2, 24]. Later on, this became a very active research field [10]; a survey can be found at [40].

Group Betweenness Centrality.

Group betweenness centrality indices [19] measure the importance of groups of nodes in networks, i.e., they measure the percentage of shortest paths that pass through at least one of the nodes of the group. On the other hand, co-betweenness centrality [32] measures the percentage of shortest paths that pass through all vertices of the group.

Algorithms for fast calculation of these group centrality measures have been developed [14, 44] even for diverse types of networks [35]; group (co-)betweenness or their variations have found applications in monitoring [17], in network formation [8], etc.

Sink Betweenness Centrality.

The notion of Sink Betweenness centrality [45], which is a specialization of our Sink Group Betweenness Centrality, was developed in the context of wireless sensor network to capture the position of nodes which lie in many paths leading to a specific node, i.e., the sink. So, the sink betweenness of a sensor node was correlated to the energy consumption of than node, since it had to relay a lot of messages towards the sink node.

3 THE SINK GROUP BETWEENNESS CENTRALITY

We assume a complex network $G(V, E)$ consisting of n nodes, where $V = \{v_i, 1 \leq i \leq n\}$ is the set of nodes and $E = \{(v_i, v_j), i, j \in V\}$ is the set of edges. We make no particular assumptions about the network being directed or undirected; this will be handled seamlessly by the underlying shortest-path finding algorithm. We assume that the network is unweighted, that is, neither the nodes nor the edges carry any weights; however, the former assumption will be reconsidered in Section 5.

Recall that the goal of sink group betweenness centrality is to discover which nodes lie in many paths leading towards a particular set of designated nodes. To achieve our purpose we combine the concepts of Group Betweenness ($\mathcal{GBC}$) [32] (although in a different fashion than in the initial definition) and the concept of Sink Betweenness ($\mathcal{SBC}$) [45]. In the following we will define the measure of Sink Group Node Betweenness Centrality ($\mathcal{SGBC}$), but firstly we will remind to the reader some definitions.

Recall that the Shortest Path Betweenness ($\mathcal{SPBC}$) Centrality is defined as follows[1]:

Definition 3.1 ((Shortest Path) Betweenness Centrality [23]). The (Shortest Path) Betweenness Centrality ($\mathcal{SPBC}$) of a node v is the

[1]When using the term "betweenness centrality" we refer to the concept of node betweenness centrality. When we wish to refer to the "edge betweenness centrality" we will explicitly make use of this term.

fraction of shortest paths between any pair of nodes that v lies on. Equation 1 calculates the $\mathcal{SPBC}$ of node v.

$$SPBC(v) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j \neq v}}^{n} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \tag{1}$$

where σ_{ij} is the number of shortest paths from node i to j, and $\sigma_{ij}(v)$ is the number of shortest paths from i to j where node v lies on. This is the non-normalized version of betweenness centrality, i.e., we do not divide by the total number of node pairs.

Definition 3.2 (Sink Betweenness Centrality [45]). The Sink Betweenness Centrality ($\mathcal{SBC}$) of a node v is the fraction of shortest paths leading to a *specific* sink node s, that v lies on. Formally, we provide Equation 2 for calculating the $\mathcal{SBC}$ of node v.

$$SBC(v) = \sum_{\substack{i=1 \\ i \neq s \neq v \\ s \text{ is the sink node}}}^{n} \frac{\sigma_{is}(v)}{\sigma_{is}} \tag{2}$$

Apparently, $\mathcal{SBC}$ is a specilization of $\mathcal{SPBC}$ i.e., j does not iterate over all nodes of the network but it is kept fixed and coincides with the sink node s ($j \equiv s$).

Suppose now that there is some “abstract grouping" process that defines non-overlapping clusters of nodes over this network. In principle, we do not need to apply any grouping algorithm at all, but we can assume that some application selects the members of each cluster as part of our input. The union of these clusters may not comprise the whole complex network. We expect that Usually the size of the aggregation of all these clusters comprises a small fraction of the complex network. Let us assume that we have defined z non-overlapping clusters, namely $C_1, C_2, \ldots, C_z$ with cardinalities $|C_1|, |C_2|, \ldots, |C_z|$, respectively. Then, the *Sink Group Betweenness Centrality* is defined as follows:

Definition 3.3 (Sink Group Betweenness Centrality). The Sink Group Betweenness Centrality ($\mathcal{SGBC}$) of a node v is the fraction of shortest paths leading to any node, which is a member of any designated cluster, that v lies on. Formally, we provide Equation 3 for calculating the $\mathcal{SGBC}$ of node v.

$$SGBC(v) = \sum_{i=1}^{n} \sum_{\substack{j \in \cup_{k=1}^{z} C_k \\ v \notin \cup_{k=1}^{z} C_k}} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \tag{3}$$

where n is the number of nodes, C_k is the k-th cluster, σ_{ij} is the number of shortest paths from node i to j, and $\sigma_{ij}(v)$ is the number of shortest paths from i to j where node v lies on.

Intuitively, a node has high Sink Group Betweenness Centrality if it sits in many shortest paths leading towards nodes belonging to any group.

Definition 3.3 requires that the node whose $\mathcal{SGBC}$ we calculate is not part of any existing cluster. This is not mandatory in general, but since we are looking for nodes which can act as mediators in the communication with the clusters' nodes, it makes sense to exclude the clusters' nodes from being considered as potential mediators. By removing this constraint, we get the concept of *Generalized* $\mathcal{SGBC}$, but in this article we consider it to be equivalent to the plain $\mathcal{SGBC}$.

We have considered unweighted and undirected complex networks. The extension of the definition of Sink Group Betweenness Centrality for directed networks is straightforward, because it is handled by the path finding algorithm. On the other hand, the extension to node or edge weighted networks is less straightforward and we will provide some ideas in a later section.

3.1 $\mathcal{SGBC}$ versus its closest relatives

$\mathcal{SGBC}$ has as its special cases the concepts of betweenness centrality and of sink betweenness centrality. In particular, $\mathcal{SGBC}$ is related to its closest relatives as follows:

- Clearly, $\mathcal{SGBC}$ is related to the $\mathcal{SPBC}$ in the following way: when the union of all clusters comprises the whole network, then *Generalized* $\mathcal{SGBC}$ and $\mathcal{SPBC}$ coincide.
- Moreover, $\mathcal{SGBC}$ is a generalization of $\mathcal{SBC}$ in the following way: when we have only one cluster which contains a single node, then $\mathcal{SGBC}$ and $\mathcal{SBC}$ coincide.
- At this point we need to make clear the difference between Group Betweenness Centrality [20] and $\mathcal{SGBC}$; the former seeks for the centrality of a group of nodes with respect to the rest of the nodes of the network, whereas the latter seeks for the centrality of a single node with respect – not the nodes of the whole network but – to the nodes belonging to some groups (clusters). Apparently, *we can generalize Definition 3.3 and Equation 3 to follow the ideas of [20].*

4 EXEMPLIFYING THE MERIT OF $\mathcal{SGBC}$

Let us now look at the small complex network of Figure 1. This network represents a collection of communicating nodes administered by an overseeing authority. We have no weights on links, but we have denoted the nodes that are mostly important for the authority – and thus must be protected better – with red color. We have not designated any attacker or attacked node in the figure, because in many situations the attack might be known where it will initiate.

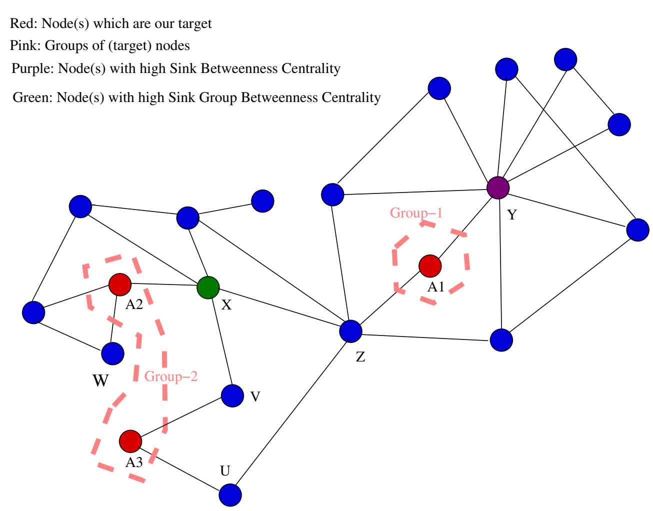

Figure 1: Illustration of Sink Group Betweenness Centrality.

Suppose that this authority is interested in *investing a limited amount of money to buy hardware/software to equip some nodes with* hoping that these will protect the significant ones, e.g., by stopping any cascade of infections. Let us call such nodes *safeguarding nodes*. Moreover, it is obvious that the authority needs to identify a limited number of safeguarding nodes, due to the limited budget. So, how we identify safeguarding nodes by exploiting the topological characteristics of the complex network?

The obvious solution is to look at the one-hop neighbors of the safeguarding nodes. Then, we can make use of the $\mathcal{SBC}$ [45] and say that the purple node (Y) is a safeguarding node. The purple node sits on many shortest paths towards the red node $A1$ (Group-1). In this way, we can select the set of nodes $\{V, W, Y\}$ as safeguarding nodes.

When the constraint of reducing the cardinality of this set comes into play, then we must somehow group safeguarding nodes, and seek for safeguarding nodes which are close to each group or to many groups. (This tradeoff with be analysed in the sequel.)

Concerning the structure of groups, we have the following characteristics:

- Clusters might contain only one node.
- Nodes comprising a cluster need not be one-hop neighbors.

$\mathcal{SBC}$ is of no use anymore, but we can use the concept of $\mathcal{SGBC}$ defined earlier. Using this concept, the green node X has high Sink Group Betweenness Centrality because it sits on many shortest paths towards the two red nodes $A2$ and $A3$ comprising Group-2. Notice, here that the nodes with high $\mathcal{SGBC}$ are not correlated to nodes with high $\mathcal{SPBC}$; for instance, the node with the highest $\mathcal{SPBC}$ in the network is node Z (it is an articulation point or bridge node).

Now let us look at the impact of safeguarding nodes' grouping on the existence (and/or $\mathcal{SGBC}$ value) of safeguarding nodes. As said earlier, the grouping creates the following tradeoff: the larger the groups we define, the less the nodes (if any) with high Sink Group Betweenness Centrality we can find.

If we unite Group-1 and Group-2 into a single group and ask the question 'which node(s) sit in many shortest paths towards ALL members of this new group', then we can safely respond that only node Z is such a node, because of the particular structure of the network of Figure 1. Recall that node Z is a bridge node, thus all shortest paths from the right of Z to nodes $A2$ and $A3$ will pass via Z, and all shortest paths from the left of Z towards node $A1$ will pass via Z.

If we now examine Figure 2, and consider initially a grouping comprised by two groups, i.e., Group-1 and Group-2, then we can clearly identify the two green nodes as those with the highest $\mathcal{SGBC}$. If we create a single large group comprised by all red nodes, then we can not find any node with high enough $\mathcal{SGBC}$ to act as safeguarding nodes. For this grouping, the blue nodes have also $\mathcal{SGBC}$ comparable to that of green nodes. Thus, with this grouping the identification of safeguarding nodes becomes problematic.

4.1 Calculation of $\mathcal{SGBC}$

The calculation of $\mathcal{SGBC}$ can be carried out using as basis the Dijkstra's algorithm. Algorithm 1 is a baseline one:

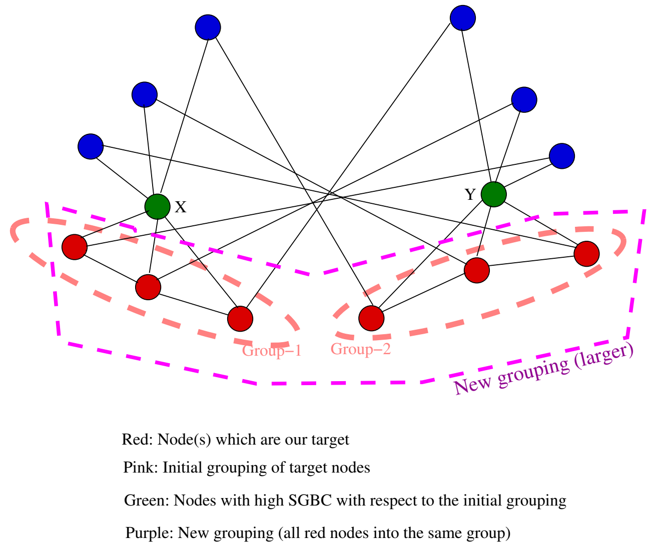

Figure 2: Impact of grouping on Sink Group Betweenness Centrality.

Algorithm 1: Calculation of $\mathcal{SGBC}$.

1 **for** *each node* $s \in \cup_{k=1}^{z} C_k$ **to** $j \in \{V - \cup_{k=1}^{z} C_k\}$ **do**
2   SP_s = Dijkstra to find shortest-path from $s \longrightarrow j$;
3 $SP = \cup_{\forall s} SP_s$;
4 **for** *each path* $p \in SP$ **do**
5   Use a hash table to group paths based on start-end;
6 **for** *each hash table bucket* **do**
7   Use a hash table to accumulate node appearance in paths;

Proposition 4.1. *The worst-case computational complexity of Algorithm 1 is* $O(|\cup_{k=1}^{z} C_k| \times n^2)$*, where n is the number of network nodes.*

Proof. Assuming a network with n nodes and m edges, and an implementation[2] of Dijkstra's algorithm that costs $O(n^2 + m)$, i.e., $O(n^2)$, then Lines 1–2 cost $O(|\cup_{k=1}^{z} C_k| \times n^2)$; lines 3–5 cost $O(n^2)$ since there are at most n^2 paths, and lines 6–7 cost $O(n^2)$. □

For unweighted and undirected networks, we can design a faster algorithm along the ideas of breadth-first traversal and the algorithm by Brandes [12], but this is beyond the scope of the present article.

5 $\mathcal{SGBC}$ WITH NODE WEIGHTS

Now suppose that each node has a weight associated with it (e.g., depicting its trustworthiness, its balance in a transaction network, etc), then we need to define the $\mathcal{SGBC}$ in such a way that it will take into account these weights. Note that this is different from having weights in edges (i.e., a weighted network) because that

[2]There are various implementations with even smaller cost utilizing sophisticated data structures or taking advantage of network sparsity.

weights are handled by the shortest-path finding algorithm. We can have the following options:

- We can apply the straightforward idea of multiplying by a node's weight as in [1], i.e.,

$$SGBC^{nw}(v) = node\_weight(v) \times \sum_{i=1}^{n} \sum_{\substack{j \in \cup_{k=1}^{z} C_k \\ v \notin \cup_{k=1}^{z} C_k}} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \quad (4)$$

  with appropriate normalization in the interval $[0 \dots 1]$ for both node's weight and $\mathcal{SGBC}$.
- Apply standard procedures for turning the problem of finding shortest paths in networks with weights on edges and/or nodes into a shortest path finding problems with weights only on edges. The methods are the following (n - number of nodes, m - number of edges):
  - We can split each node apart into two nodes as follows: for any node u, make two new nodes, u_1 and u_2. All edges that previously entered node u now enter node u_1, and all edges that leave node u now leave u_2. Then, put an edge between u_1 and u_2 whose cost is the cost node u_1. In this new graph, the cost of a path from one node to another corresponds to the cost of the original path in the original graph, since all edge weights are still paid and all node weights are now paid using the newly-inserted edges. Constructing this graph can be in time $O(m + n)$, since we need to change each edge exactly once and each node exactly once. From there, we can just use a normal Dijkstra's algorithm to solve the problem in time $O(m+nlogn)$, giving an overall complexity of $O(m + nlogn)$. If negative weights exist, then we can use the Bellman-Ford algorithm instead, giving a total complexity of $O(mn)$.
  - Alternatively, we can think as follows: since we have both edges and nodes weighted, when we move from i to j, we know that total weight to move from i to j is weight of edge($i \rightarrow j$) plus weight of j itself, so lets make $i \rightarrow j$ edge weight sum of these two weights and the weight of j zero. Then we can find shortest path from any node to any other node to in $O(mlogn)$.
- We can multiply each fraction (of shortest paths) in the summation formula with the minimum weight (positive or negative) found along the path.

6 SINK GROUP EDGE BETWEENNESS CENTRALITY

The plain edge-betweenness centrality measure is used to identify the edges which lie in many shortest-paths among pair of network nodes. It has been widely used in network science and not only, e.g., for discovering communities in networks [25], for designing topology control algorithms for ad hoc networks [15], etc.

In our context, we ask the question whether we can identify the edges which lie in many paths towards a set of subsets of network nodes. The extension of $\mathcal{SGBC}$ to the case of edges is easy. Thus, the *Sink Group Edge Betweenness Centrality* is defined as follows:

Definition 6.1 (Sink Group Edge Betweenness Centrality). The Sink Group Edge Betweenness Centrality ($\mathcal{SGEBC}$) of an edge e is the fraction of shortest paths leading to any node, which is a member of any designated cluster, that e lies on. Formally, we provide Equation 5 for calculating the $\mathcal{SGEBC}$ of edge e.

$$SGEBC(e) = \sum_{i=1}^{n} \sum_{\substack{j \in \cup_{k=1}^{z} C_k \\ v \notin \cup_{k=1}^{z} C_k}} \frac{\sigma_{ij}(e)}{\sigma_{ij}} \quad (5)$$

where n is the number of nodes, C_k is the k-th cluster, σ_{ij} is the number of shortest paths from node i to j, and $\sigma_{ij}(e)$ is the number of shortest paths from i to j where edge e lies on.

7 APPLICATIONS OF $\mathcal{SGBC}$

7.1 $\mathcal{SGBC}$ and influence minimization

We have already explained in the introduction how sink group betweenness centrality can be used to limit the spreading in the context of influence or infection minimization problems under budget constraints. Especially relevant becomes for those problems that are online [7], i.e., require a continuous combat against the spreading while the infection evolves in various parts of the network.

7.2 $\mathcal{SGBC}$ and virtual currencies networks

Let us now look again at the network of Figure 1, but this time assume that the graph represents a network of transactions being made using a virtual currency, e.g., a community currency [46], which – differently from BitCoin – is administered by an overseeing authority. We have denoted the nodes in deficit (i.e., with lack of virtual money) with red color.

Suppose that this authority is interested in injecting a limited amount of money into some nodes with the hope that these nodes will buy something from the red nodes and therefore with reduce their deficit. Let us call such nodes *deficit balancers*. Suppose that the authority needs to identify a limited number of deficit balancers, otherwise will have to divide this amount of money into too many deficit balancers, and eventually an even smaller amount of money will end up to the nodes with deficit. So, how we identify deficit balancers?

Exactly as before, the obvious solution is to look at the one-hop neighbors of the deficited nodes. Then, we can make use of the $\mathcal{SBC}$ and say that the purple node (Y) is a deficit balancer. The purple node sits on many shortest paths towards the red node $A1$ (Group-1). In this way, we can select the set of nodes $\{V, W, Y\}$ as deficit balancers. If the authority is constrained to reduce the cardinality of this set, then we must seek for deficit balancers which are close to each group or to many groups. From this point, we can continue along the same reasoning as we did in Section 4.

7.3 $\mathcal{SGBC}$ and community finding

The concept of $\mathcal{SGEBC}$ can be used as a component of an algorithm for discovering the community where a particular set of nodes belongs. The idea is that if this collection of nodes is relative close to each other, then repeated deletion of high $\mathcal{SGEBC}$ edges will gradually isolate this community of nodes from the rest of the network nodes.

8 CONCLUSIONS

Network science continues to be a very fertile research and development area. The more our world becomes connected via 5G and Internet of Things (or Everything), the more network science evolves into a precious tool for graph-data analysis. One of the central concepts in this field, namely centralities despite counting half a century of life, is still hot giving new definitions, and new insights into the networks' organization. In this article, we introduced a new member into the family of shortest path betweenness centralities, namely sink group betweenness centrality. The purpose of this measure is to identify nodes which are in positions to monitor/control/mediate the communication towards subsets of networks nodes. We provided a simple algorithm to calculate this measure, and also extended it for node-weighted networks, and also for the case of edge betweenness. This effort is simply out first step in a long journey to develop efficient algorithms for the calculation of these measures, to analyze its distribution in real networks, to extend them to consider sink group betweenness centrality computed for a collections of network nodes, and develop techniques for approximating them.

ACKNOWLEDGMENTS

Yannis Manolopoulos would like to thank the Aristotle University of Thessaloniki, Greece, where he serves as Professor Emeritus.

REFERENCES

[1] G. A. A. Akanmu, F. Z. Wang, and H. Chen. 2012. Introducing weighted nodes to evaluate the cloud computing topology. *Journal of Software Engineering & Applications* 5, 11 (2012), 961–969.
[2] D. A. Bader, S. Kintali, K. Madduri, and M. Mihail. 2007. Approximating betweenness centrality. In *Proceedings of the International Workshop on Algorithms & Models for the Web-Graph (WAW)*. 124–137.
[3] A. L. Barabasi. 2016. *Network Science*. Cambridge University Press.
[4] A. L. Barabasi and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286, 5439 (1999), 509–512.
[5] P. Basaras, G. Iosifidis, D. Katsaros, and L. Tassiulas. 2019. Identifying influential spreaders in complex multilayer networks: A centrality perspective. *IEEE Transactions on Network Science & Engineering* 6, 1 (2019), 31–45.
[6] P. Basaras, D. Katsaros, and L. Tassiulas. 2013. Detecting influential spreaders in complex, dynamic networks. *IEEE Computer Magazine* 46, 4 (2013), 26–31.
[7] P. Basaras, D. Katsaros, and L. Tassiulas. 2015. Dynamically blocking contagions in complex networks by cutting vital connections. In *Proceedings of the IEEE International Conference on Communications (ICC)*. 1170–1175.
[8] X. Bei, W. Chen, S.-H. Teng, J. Zhang, and J. Zhu. 2011. Bounded budget betweenness centrality game for strategic network formations. *Theoretical Computer Science* 412 (2011), 7147–7168.
[9] G. Bianconi. 2018. *Multilayer Networks: Structure and Function*. Oxford University Press.
[10] M. G. Bonchi, F. De Francisci and Riondato M. 2016. Centrality measures on big graphs: Exact, approximated, and distributed algorithms. In *Proceedings of the ACM International Conference on the World Wide Web (WWW)*. 1017–1020.
[11] J. Borge-Holthoefer and Y. Moreno. 2012. Absence of influential spreaders in rumor dynamics. *Physical Review E* 85, 2 (2012).
[12] U. Brandes. 2001. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology* 25, 2 (2001), 163–177.
[13] U. Brandes. 2008. On variants of shortest-path betweenness centrality and their generic computation. *Social Networks* 30, 2 (2008), 136–145.
[14] M. H. Chehreghani. 2014. Effective co-betweenness centrality computation. In *Proceedings of the ACM Conference on Web Search & Data Mining (SDM)*. 423–432.
[15] A. Cuzzocrea, A. Papadimitriou, D. Katsaros, and Y. Manolopoulos. 2012. Edge betweenness centrality: A novel algorithm for QoS-based topology control over wireless sensor networks. *Journal of Network & Computer Applications* 35, 4 (2012), 1210–1217.
[16] S. Dolev, Y. Elovici, and R. Puzis. 2010. Routing betweenness centrality. *J. ACM* 57, 4 (2010).
[17] S. Dolev, Y. Elovici, R. Puzis, and P. Zilberman. 2009. Incremental deployment of network monitors based on Group Betweenness Centrality. *Inform. Process. Lett.* 109 (2009), 1172–1176.
[18] D. Easley and J. Kleinberg. 2010. *Networks, Crowds and Markets: Reasoning About a Highly Connected World*. Cambridge University Press.
[19] M. G. Everett and S. P. Borgatti. 1999. The centrality of groups and classes. *Journal of Mathematical Sociology* 23, 3 (1999), 181–201.
[20] M. G. Everett and S. P. Borgatti. 2009. The centrality of groups and classes. *The Journal of Mathematical Sociology* 23, 3 (2009), 181–201.
[21] S. Fortunato. 2010. Community detection in graphs. *Physics Reports* 486, 3-5 (2010), 75–174.
[22] S. Fortunato and D. Hric. 2016. Community detection in networks: A user guide. *Physics Reports* 659 (2016), 1–44.
[23] L. C. Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry* 40, 1 (1977), 35–41.
[24] R. Geisberger, P. Sanders, and D. Schultes. 2008. Better approximation of betweenness centrality. In *Proceedings of the Meeting on Algorithm Engineering & Experiments (ALENEX)*. 90–100.
[25] M. Girvan and M. E. J. Newman. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99 (2002), 7821–7826.
[26] A. Halu, R. J. Mondragón, P. Panzarasa, and G. Bianconi. 2013. Multiplex PageRank. *PLOS One* 8, 10 (2013), e78293.
[27] X. Huang, D. Chen, T. Ren, and D. Wang. 2021. A survey of community detection methods in multilayer networks. *Data Mining & Knowledge Discovery* 35 (2021), 1–45.
[28] W. Hwang, T. Kim, M. Ramanathan, and A. Zhang. 2008. Bridging centrality: Graph mining from element level to group level. In *Proceedings of the ACM Conference on Knowledge Discovery & Data Mining (KDD)*. 336–344.
[29] D. Katsaros and P. Basaras. 2015. Detecting influential nodes in complex networks with range probabilistic control centrality. In *Coordination Control of Distributed Systems*, J. H. van Schuppen and T. Villa (Eds.). Lecture Notes in Control and Information Sciences, Vol. 456. Springer-Verlag, 265–272.
[30] D. Katsaros, N. Dimokas, and L. Tassiulas. 2010. Social network analysis concepts in the design of wireless ad hoc network. *IEEE Network Magazine* 24, 6 (2010).
[31] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse. 2010. Identification of influential spreaders in complex networks. *Nature Physics* 6 (2010), 888–893.
[32] E. D. Kolaczyk, D. B. Chua, and M. Barthelemy. 2009. Group betweenness and co-betweenness: Inter-related notions of coalition centrality. *Social Networks* 31, 3 (2009), 190–203.
[33] A. N. Langville and C. D. Meyer. 2006. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press.
[34] V. Lattora, V. Nicosia, and G. Russo. 2017. *Complex Networks: Principles, Methods and Applications*. Cambridge University Press.
[35] L. Li, G. Wang, and M. Yu. 2015. Pairwise co-betweenness for several types of network. *Journal of Networks* 10, 2 (2015), 91–98.
[36] L. Maccari, L. Ghiro, A. Guerrieri, A. Montresor, and R. Lo Cigno. 2020. Exact distributed load centrality computation: Algorithms, convergence, and applications to distance vector routing. *IEEE Transactions on Parallel & Distributed Systems* 31, 7 (2020), 1693–1706.
[37] N. Magaia, A. P. Francisco, P. Pereira, and M. Correia. 2015. Betweenness centrality in Delay Tolerant Networks: A survey. *Ad Hoc Networks* 33 (2015), 284–305.
[38] P. Mahendra, P. Mikhail, and H. Liaquat. 2013. Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks. *PLOS One* 8, 1 (2013).
[39] N. Masuda and R. Lambiotte. 2016. *A Guide to Temporal Networks*. Series on Complexity Science, Vol. 4. World Scientific.
[40] J. Matta, G. Ercal, and K. Sinha. 2019. Comparing the speed and accuracy of approaches to betweenness centrality approximation. *Computational Social Networks* 6 (2019).
[41] M. E. J. Newman. 2010. *Networks: An Introduction*. Oxford University Press.
[42] P. Pantazopoulos, M. Karaliopoulos, and I. Stavrakakis. 2014. Distributed placement of autonomic Internet services. *IEEE Transactions on Parallel & Distributed Systems* 25, 7 (2014), 1702–1712.
[43] D. Papakostas, T. Kasidakis, E. Fragkou, and D. Katsaros. 2021. Backbones for Internet of Battlefield Things. In *Proceedings of the IEEE/IFIP Conference on Wireless On-demand Network Systems & Services (WONS)*. 116–123.
[44] R. Puzis, Y. Elovici, and S. Dolev. 2007. Fast algorithm for successive computation of group betweenness centrality. *Physical Review E* 76 (2007), 056709.
[45] H. S. Ramos, A. C. Frery, A. Boukerche, E. M. R. Oliveira, and A. A. F. Loureiro. 2014. Topology-related metrics and applications for the design and operation of wireless sensor networks. *ACM Transactions on Sensor Networks* 10, 3 (2014), 1–35.
[46] G. Seyfang and N. N. Longhurst. 2013. Growing green money? Mapping community currencies for sustainable development. *Ecological Economics* 86 (2013), 65–77.