Network Science **7** (3): 376–401, 2019
doi:10.1017/nws.2019.7

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Centralities for networks with consumable resources

Hayato Ushijima-Mwesigwa[1*], Zadid Khan[2], Mashrur A. Chowdhury[2] and Ilya Safro[3]

[1]School of Computing, Clemson University, Clemson, SC, USA, [2]Department of Civil Engineering, Clemson University, Clemson, SC, USA (emails: mdzadik@clemson.edu, mac@clemson.edu), [3]School of Computing, Clemson University, Clemson, SC, USA (email: isafro@clemson.edu)
*Corresponding author. Email: hushiji@g.clemson.edu

Action Editor: Dr. Filippo Menczer

Abstract

Identification of influential nodes is an important step in understanding and controlling the dynamics of information, traffic, and spreading processes in networks. As a result, a number of centrality measures have been proposed and used across different application domains. At the heart of many of these measures lies an assumption describing the manner in which traffic (of information, social actors, particles, etc.) flows through the network. For example, some measures only count shortest paths while others consider random walks. This paper considers a spreading process in which a resource necessary for transit is partially consumed along the way while being refilled at special nodes on the network. Examples include fuel consumption of vehicles together with refueling stations, information loss during dissemination with error-correcting nodes, and consumption of ammunition of military troops while moving. We propose generalizations of the well-known measures of betweenness, random-walk betweenness, and Katz centralities to take such a spreading process with consumable resources into account. In order to validate the results, experiments on real-world networks are carried out by developing simulations based on well-known models such as Susceptible-Infected-Recovered and congestion with respect to particle hopping from vehicular flow theory. The simulation-based models are shown to be highly correlated with the proposed centrality measures.

Reproducibility: Our code and experiments are available at `https://github.com/hmwesigwa/soc_centrality`

Keywords: network centrality, Katz centrality, betweenness centrality, random-walk betweenness centrality, consumable resources

Author ORCID. Hayato Ushijima-Mwesigwa, https://orcid.org/0000-0002-1442-3077

1. Introduction

Spreading processes are ubiquitous throughout science, nature, and society (Strogatz, 2001; Tao et al., 2006; Pastor-Satorras & Vespignani, 2001). These include spreading of infectious diseases (Keeling & Rohani, 2011), computer viruses (Kephart et al., 1997), cascading failures (Motter, 2004), traffic congestion (Li et al., 2015), opinion spreading (Liu et al., 2007; Bettencourt et al., 2006), and reaction-diffusion processes (Colizza et al., 2007). Understanding a nodes' spreading influence is fundamental for a wide variety of applications such as epidemiology (Diekmann & Heesterbeek, 2000; Keeling & Rohani, 2011), viral marketing (Watts et al., 2007; Leskovec et al., 2007a), collective dynamics (Albert & Barabási, 2002; Boccaletti et al., 2006; Barrat et al., 2008) and robustness of networks (Albert et al., 2000; Newman, 2003; Cohen et al., 2001), and so forth. Whereas many centrality measures were originally developed for social networks,

© Cambridge University Press 2019

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

some of them have subsequently been adapted to quantify the importance of nodes in epidemiological spreading processes (Kitsak et al., 2010; Ghosh & Lerman, 2012; Šikić et al., 2013; Bae & Kim, 2014). This is partly due to the fact that most centralilty measures have simple assumptions, thus these measures are often intuitive and interpretable for a given application. Moreover, most popular centrality measures are based on variants of paths and eigenvector computations which explain paradigms in spreading models. Popular centrality measures include degree, closeness, betweenness, current-flow, PageRank, eigenvector, and Katz centralities (Freeman, 1978; Newman, 2005; Brandes & Fleischer, 2005; Page et al., 1999; Bonacich, 1987, 1991; Katz, 1953). All these measures make an implicit assumption about the process in which a commodity (e.g., information, vehicles, or infection) flows in the network. Typically, closeness and betweenness assume flow on geodesic paths, while PageRank, eigenvector, and Katz centrality model flow via random walks. The extent to which a centrality measure can be interpreted for a given application depends on whether or not the assumed flow characteristics are a good representation of what is actually flowing in the network.

In this work, we consider a flow process in which a resource essential for flow is consumed along the way and can be refilled at specially assigned nodes in order to ensure that a flow process is not terminated. For example, a vehicle consumes fuel as it travels in a network that has refueling station nodes. In another domain, information requires updating or refreshing while it moves over the network. For example, in real information and social networks, rumors and gossips often die out if not refreshed (Moreno et al., 2004) and forgetting rates are considered in models (Zhao et al., 2013). Not much of the existing work models well-known concepts on networks by taking into account consumable resources.

For simplicity, we model the resource consumption as a discrete process that limits the number of steps the flow process can take without refilling the resource. For a graph underlying network of interest, $G = (V, E)$, the parameter κ represents the number of steps a process can take without a consumed resource being refilled, and $\Omega \subset V$ represents the refilling nodes. One of the most intuitive and important modern applications of this process is the in-motion recharging of electric vehicle batteries that is anticipated to be broadly implemented in future. We borrow terminology from the charge level of batteries for electric vehicles and refer to the currently available resource as the *state of charge* (SOC). Thus, κ represents the full SOC value. In the next section, a list of related real-word applications is given.

1.1 Our contribution

In this work, we study a process where a commodity (such as information and traffic) is flowing (or spreading) in a network while consuming a resource necessary for flow and being refilled at special nodes. We give a list of potential applications that have a similar flow (or spreading) process. In order to estimate a nodes' spreading influence, we generalize the measures of Katz, betweenness, and random-walk betweenness centralities (RWBC) (including its generalization for directed graphs) and show how they can be computed. Lastly, we present different models to simulate the spreading processes and show that the generalized centrality measures are highly correlated with the simulation-based models.

2. Applications

Transportation networks: A natural and motivating application of such a process is in transportation networks, in particular, road networks. Let a node in a road network represent a road segment. An edge between two road segments exists if they are physically adjacent to each other (as prolongation of each other or with a real intersection). In this case, $\Omega \subset V$ represents the nodes with refueling stations, and κ - the maximum distance a vehicle can travel without refueling. In particular, we can also consider electric vehicle road networks equipped with wireless charging lanes, where a whole lane can be turned into a charging infrastructure. This technology has

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

seen tremendous growth over the last couple of years with test sites already in place (Jang et al., 2012). However, setting up this technology will come with a heavy price tag for a city with a limited budget, thus tools must be developed to analyze these networks beforehand. While several research studies have carried out for identifying the optimal locations for the deployment of wireless charging lanes (Riemann et al., 2015; Chen et al., 2016; Ushijima-Mwesigwa et al., 2017; Khan et al., 2018, 2019), these studies often make different assumptions while solving different optimization objectives and constraints. Having an independent tool to analyze the network is thus also necessary in order for deployment strategies to be compared.

Peer-to-peer networks: Peer-to-peer (P2P) information exchange systems (Basu et al., 2013) have gained popularity over the past two decades. One of the challenges in P2P systems is searching for content on the network. Gnutella is a popular open and decentralized file-sharing protocol in P2P networks (Wang et al., 2007). The Gnutella protocol works as follows (Ripeanu & Foster, 2002):

1. A node (computer) v connects to the Gnutella network by connecting to a set of one or more nodes, U, already in the network. Then v announces its existence to all nodes in U.
2. The nodes in U announce to all their neighbors that v has joined the network, which also announce to their neighbors, and so forth.
3. Once all nodes are aware of v's existence, it can make a query on the network.

Popular methods of message propagation for a given query issued by a node include such methods as flood-based and random-walks routing algorithms (Tsoumakos & Roussopoulos, 2006). A global time-to-live (TTL) parameter represents the maximum number of steps (also known as hops) a query can take before it gets discarded. In a flood-based routing algorithm, a querying node contacts all its neighbors, who then contact all their neighbors, and so forth. The process stops after each message has taken TTL number of steps. This simplistic method produces a huge overhead by contacting many nodes. In the random-walks routing algorithm, the querying node randomly chooses a subset of its neighbors and sends each of them k messages, for some k. Each of these messages starts its own random walk in the network that is terminated after TTL steps. Other termination conditions exist; however, they are not relevant for this work. The random-walks routing algorithm greatly reduces the message passing throughout the network, with other advantages such as local load balancing, since no nodes are favored over others during message propagation. However, depending on the network topology, success rates could vary significantly. These two routing algorithms are sometimes referred to as blind search methods. On the other hand, informed search methods include methods that, for example, take advantage of previous queries making better decisions for message passing, and in an ideal scenario, a message could then take the shortest path to a target node.

In the Gnutella network, once a node receives a message, it first reduces the TTL counter of the message before forwarding it. The TTL parameter is intuitively equivalent to the SOC parameter κ in this paper while the set of nodes in Ω represents nodes that reset the TTL counter before forwarding the message. These, for example, could be compromised nodes. In the analysis of the Gnutella network, an interesting question is determining the most important nodes in the network, which could be the nodes that receive the most traffic.

Social networks (online): The popularity and complexity of online social networks (OSNs) have seen a tremendous growth in the last two decades and will continue to grow. In OSNs such as Twitter or Facebook, a user shares information which can be viewed by other users he/she is connected to. Centrality measures are often used to identify influential users within an OSN. However, in most centrality measures, all the users in the network are assumed to exhibit similar behavioral features. In other words, they all have the same desire and motivation to share knowledge in the network. In most cases, all the users are assumed to be *active* users, users that are willing and motivated to share their knowledge. Many studies have been carried out and they show that this is not the case. In fact, most users, while being beneficiaries of the content being shared, actually do

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

not share information themselves. The two different types of users are often referred to as *posters* and *lurkers* (Marett & Joshi, 2009; Lai & Chen, 2014; Preece et al., 2004; Schlosser, 2005). Posters are defined in Schlosser (2005) as the active users who share their experiences and create content on the internet while lurkers are defined as the passive users who do not necessarily create any content. The analysis of lurkers in social networks is now an active area of research (Interdonato & Tagarelli, 2016). In order to have a better understanding of OSNs, it is important to understand how information and content are shared over the network. It has been reported that lurkers are the majority in many online communities. The percentage of lurkers in an online community varies across different studies with some giving estimates as high as 90% of the total users (Van Mierlo, 2014). Sometimes referred to as the *participation inequality principle* (Interdonato & Tagarelli, 2016), in which only a small percentage of users actually contribute to the online content while the rest never do. Lurkers are not registered users who do not use their account; they can share information in subtle ways. For example, Facebook has the "like" feature and thus user's contacts can see the information he/she liked. In this work, we take the posters to be members of the set Ω and assume that a piece of information is coupled with a *momentum* or *penetrating power*. Following the analogy of a battery charge in Electric Vehicle (EV), the momentum of a piece of information has the power to drive the information for a limited number of steps, κ, before it dies out. However, if it reaches a node in Ω, it regains its momentum.

Personalized web ranking: Consider a person surfing the web at random, however, with a topic of interest in mind. The surfer begins at a web page ω_1 and performs a random walk on the web graph. At each time step, the person proceeds from the current page u to a randomly chosen web page that u is linked to. If after κ steps, the surfer has not found a web page of interest, the surfer starts the process again from ω_1. However, if the surfer discovers a web page ω_2 that is relevant to the topic of interest, ω_2 becomes the new restarting point and the process continues. If Ω is the set of web pages known beforehand, then centrality measure defined by such a process can be used to rank the web pages based on the ones in Ω, giving a personalized page ranking strategy. In this class of applications, we can also mention random-walk-based similarity measures on graphs and hypergraphs (Shaydulin et al., 2017; Fouss et al., 2007; Chen & Safro, 2011) that would benefit from introducing resource consumption restrictions for the distance of a random walk.

3. Related work

Centrality as a way of analyzing social networks dates back to Bavelas (1948). Since then, various methods of centrality have been proposed to quantify the importance of individuals in social networks. These measures have also been effectively used as tools to study networks in other diverse fields such as physics, biology, and engineering.

Since our initial motivation for this research was related to the analysis of road networks, in this section, we first highlight how previous studies have used centrality measures to analyze road networks. Next, we briefly introduce and summarize studies on electric vehicle road networks. Lastly, based on the potential applications for the proposed centrality measures, we summarize different existing approaches for possible applications.

We briefly summarize an incomplete list of studies in which centralities have been used to study and analyze road networks. A road network pattern can be viewed as the geographical layout and structure of a network. A road network can be laid out in different patterns (e.g., see the book by Southworth & Ben-Joseph (2013) for more information on road network patterns) which can affect traffic performance, travel behavior, and traffic safety. In Zhang et al. (2011), the betweenness centrality is computed to analyze and classify road network patterns. In particular, it is used to define a measure that can quantitatively distinguish between different pattern types. In Wang et al. (2012), the authors use centrality measures to analyze road networks in urban areas and apply their findings to mitigate congestion. More specifically, they use large-scale mobile phone

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

data, with detailed Geographic Information System data, to detect types of road usage and determine the origins of the drivers. This information is used to build a bipartite network with nodes representing road segments and *driver sources*, which is then called the *network of road usage*. Here, a driver source is a zone where the mobile phone user lives. This can be located using the mobile phone data. Given a list of all driver sources, for each road segment r, the authors calculated the fraction of traffic flow on r that was generated by each driver source. They then ranked the driver sources by their contribution to the traffic flow. Based on this information, an edge in the network of road usage exists between a road segment r and the top-ranked source nodes that produce 80% of r's traffic flow. Finally, the betweenness centrality of a road segment r in the road network, along with the degree centrality of r in the network of road usage, was used to classify and group the road segments in the network. Experiments carried out in the San Francisco Bay area and Boston area provide evidence to show that the findings could enable cities to tailor targeted strategies to reduce the average daily commute time.

In Scheurer et al. (2008), the authors used a wide variety of centrality measures such as betweenness, closeness, and degree centrality to identify the positive and negative points of the public transportation networks from different perspectives such as coverage, connectivity, and service levels. The collective human spatial movement behavior is explored in Jiang & Jia (2011). The authors use, among others, PageRank and betweenness centrality coupled with agent-based simulations to study the movement of pedestrians in London street network. In Altshuler et al. (2011), the authors propose an estimation method for mobility prediction in transportation networks based on the betweenness centrality carrying out experiments on the Israeli transportation network. Other studies on transportation networks, where centrality plays a crucial role in the analysis, include Jayasinghe et al. (2015), Jiang & Claramunt (2004), Porta et al. (2006), Jayaweera et al. (2017), Crucitti et al. (2006), and Park & Yilmaz (2010).

As cities move toward reducing their carbon footprint, EVs offer the potential to reduce both petroleum imports and greenhouse gas emissions. However, batteries in these vehicles have a limited travel distance per charge. This results in a major obstacle for EV widespread adaptation, namely, *range anxiety*, the persistent worry about not having enough battery power to complete a trip. The emergence of EV wireless charging technology where a whole lane can be turned into a charging infrastructure provides itself as a potential solution to range anxiety. For a more detailed study of the design, application, and future prospects of this technology, the reader is encouraged to see, for example, Qiu et al. (2013), Bi et al. (2016), Li & Mi (2015), Lukic & Pantic (2013), Cirimele et al. (2014), Fuller (2016), Vilathgamuwa & Sampath (2015), Ning et al. (2013), Yan et al. (2006), and Guimerà et al. (2002). With a heavy price tag, a deployment of this technology without a careful study can lead to inefficient use of limited resources. One of the main purposes of this paper is to provide a tool to study and analyze road networks with a given deployment of wireless charging lanes. In these EV road networks, we assume that in order for an EV to travel between any two nodes, it is possible that a vehicle may need to detour to get charged to arrive at its destination. We envision that our modified version of betweenness centrality can be used in studying these EV road networks in similar ways as the studies in the preceding paragraph.

Studies analyzing social networks whose users can be categorized as posters and lurkers have recently been gaining attention. In Tagarelli & Interdonato (2013, 2014), the authors propose centrality measures for ranking lurkers in social networks. In these works, no prior knowledge of whether a user is a lurker/poster or not is assumed. The authors define a topology-driven lurking framework to model the relationships from information-producer to information-consumer. As a result, lurkers are ranked based on only the topology of the network. In our applications of the proposed centrality measures, we assume prior knowledge of whether or not a user is a poster or lurker. The main basis for this assumption is that the network topology may not be related to a user's desire to share information. For example, two users on Facebook may have the same number of connections; however, they have very different desires to share or post information.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

4. Graph model

Let $G = (V, E)$ be an unweighted (directed or undirected) graph underlying a network of interest. The underlying assumption is that the commodity (such as information and moving vehicle) is flowing (or spreading) on G while consuming a resource necessary for flow. The flow is limited to the nodes within a geodesic distance of at most κ edges, $\kappa \in \mathbb{N}$. In addition, there exists a subset of nodes $\Omega \subset V$ that refill the resource. In other words, if the commodity passes through a node $u \in \Omega$, it can then spread further to nodes that are at most κ edges, from u. This process is modeled by a directed graph with adjacency matrix $\mathcal{B}_\kappa$ (see below).

Let A be the adjacency matrix of G, and $|V| = n$. Define the state space of a commodity traversing G as the set

$$\mathscr{V} := \{(u, i) | u \in V, 0 \leq i \leq \kappa\},$$

in which the state (u, i) represents the event that the commodity is at node $u \in V$ with the current level of SOC at i. The transition from one state to another is modeled by a directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, where $\mathscr{E} = \mathscr{E}_1 \cup \mathscr{E}_2$ with

$$\mathscr{E}_1 := \big\{\big((u, i), (v, \kappa)\big) \mid (u, v) \in E, v \in \Omega\big\},$$

and

$$\mathscr{E}_2 := \big\{\big((u, i), (v, j)\big) \mid (u, v) \in E,\ i = j + 1, v \notin \Omega\big\}.$$

The set $\mathscr{E}_1$ represents a transition where the current SOC is increased to κ (refilled), while $\mathscr{E}_2$ represents a transition where the current SOC is reduced by 1 (consumed). The adjacency matrix of $\mathscr{G}$, denoted as $\mathcal{B}_\kappa$, is defined as follows. Let J_Ω be a diagonal $n \times n$ matrix given by

$$[J_\Omega]_{ii} = \left\{ \begin{array}{ll} 1, & \text{if } i \in \Omega \\ 0, & \text{otherwise.} \end{array} \right\}$$

For ease of exposition, where it is clear, we drop the subscript in J_Ω and simply write J. If I the $n \times n$ identity matrix defines the $n(\kappa + 1) \times n(\kappa + 1)$ block matrix $\mathcal{B}_\kappa$ as

$$\mathcal{B}_\kappa = \begin{bmatrix} \mathrm{AJ} & \mathrm{A(I\text{-}J)} & \mathbf{0} & \dots & \dots & \dots & \mathbf{0} \\ \mathrm{AJ} & \mathbf{0} & \mathrm{A(I\text{-}J)} & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \mathrm{AJ} & \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & & & \ddots & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ \mathrm{AJ} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \dots \mathrm{A(I\text{-}J)} & \\ \mathrm{AJ} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \tag{1}$$

then the block matrix $\mathcal{B}_\kappa$ defines a directed state space graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ that models the underlying flow process. In order for a commodity to flow from node s to t, it may be necessary to traverse one or more nodes in Ω. With this in mind, we define a feasible walk to represent the walks in G that the commodity can fully traverse.

Definition 1. *A walk* w *in* G *is a* feasible walk *if a commodity starting with full SOC can traverse w.*

The proposed centrality measures in the following sections are based on computing the feasible walks in the network.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

5. Katz centrality

The centrality measure proposed by Katz (1953) was originally intended to rank a node (i.e., an actor in a social system) influence within a social network according to the number of its contacts considering different path lengths to other nodes. Thus, the model takes into account not only the immediate neighbors of a node but also its neighbors of second-order, third-order, and so on. The computation of Katz centrality is based on random walks emanating from a node. In this section, we propose SOC-Katz centrality that only takes feasible walks into account.

5.1 Counting number of feasible walks

For an adjacency matrix A of a graph, the ijth entry of the matrix A^k, $k \in \mathbb{N}$, counts paths from i to j of length k. *However, in our resource consumption model, not all of these walks are in fact feasible.* This leads to an interesting question of finding a matrix that represents the number of i-j feasible walks.

Consider the matrix $\mathcal{B}_\kappa^k$. Assume that the index of nodes in V and matrices $I, J, A, \mathcal{B}_\kappa$ start at 0. For $i, j \in V$, with $0 \le i < n$, $0 \le i' < n(\kappa+1)$, and $j \equiv i' \pmod{n}$, the ijth entry of $\mathcal{B}_\kappa^k$ gives the number of walks from node i to j completing with a different SOC. The destination SOC is given by the value $\lfloor i'/n \rfloor$. This implies that the number of walks from i to j ending with non-negative SOC is given by the summation at each SOC level.

Let $\mathcal{I}_\kappa$ be an $n(\kappa+1) \times n$ block matrix with $\kappa+1$ blocks of identity matrix I given by

$$\mathcal{I}_\kappa = \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix}, \text{ and } \mathcal{Z}_\kappa = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2}$$

Then the matrix $\mathcal{I}_\kappa^T \mathcal{B}_\kappa^k \mathcal{I}_\kappa$ is an $n \times n$ matrix whose ijth entry gives the number of feasible walks from i to j of length k. Let S be the $n(\kappa+1) \times n(\kappa+1)$ matrix with ij term given by

$$s_{ij} = \sum_{k=1}^{\infty} \alpha^k [\mathcal{B}_\kappa^k]_{ij} \tag{3}$$

Thus,

$$\begin{aligned} S &= I_{n(\kappa+1)\times n(\kappa+1)} + \alpha\mathcal{B}_\kappa + \alpha^2\mathcal{B}_\kappa^2 + \cdots + \alpha^i\mathcal{B}_\kappa^i + \cdots \\ &= (I_{n(\kappa+1)\times n(\kappa+1)} - \alpha\mathcal{B}_\kappa)^{-1} \end{aligned}$$

Then, if W is the $n \times n$ matrix given

$$W = \mathcal{Z}_\kappa^T (I_{n(\kappa+1)\times n(\kappa+1)} - \alpha\mathcal{B}_\kappa)^{-1} \mathcal{I}_\kappa \tag{4}$$

The centrality is then given by $C = W\mathbf{1}$, where $\mathbf{1}$ is the column vector consisting of all 1's. For the standard centrality measure, the Katz centrality is computed by $(I - \alpha A)^{-1}$. The parameter α, also known as the *damping factor*, must be chosen carefully such that $0 < \alpha < 1/\lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of $\mathcal{B}_\kappa$.

Lemma 5.1. $1/\lambda_{\max}(A) \le 1/\lambda_{\max}(\mathcal{B}_\kappa)$.

Proof. If $\mathcal{A} = \mathcal{Z}_m^T \mathcal{B}_m \mathcal{I}_m$, then $\mathcal{A}^k$ is the $n \times n$ matrix that counts the walks of length k in the bounded-walk graph. Clearly, $[A^k]_{ij} \ge [\mathcal{A}^k]_{ij}$ for all i, j. This implies that if the sequence $\{\alpha^k A^k\}_{k=1}^{\infty}$ converges, then $\{\alpha^k \mathcal{A}^k\}_{k=1}^{\infty}$ converges. Thus, $\lambda_{\max}(A) > \lambda_{\max}(\mathcal{B})$ □

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

When $\alpha \to 0$, then only walks of very short length are taken into account and degree centrality usually performs well (Borgatti, 2005; Kitsak et al., 2010). However, as the value of α increases, eigenvector and Katz outperform other measures (Liu et al., 2016). Due to lemma 5.1, we are able to take larger values of α compared to the standard Katz centrality measure.

6. Betweenness centrality

For a graph G, let σ_{st} be the total number of shortest paths from nodes s to t, while $\sigma_{st}(v)$ be the total number of shortest paths from s to t that pass through v. Then the (unnormalized) betweenness centrality of v, $BC(v)$, is given by

$$BC(v) := \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{5}$$

The decision of whether or not to include the end-points of a path to fall on that path is usually made according to specific applications and goals. This is because the only difference this makes is an additive constant to $BC(v)$. In this paper, we will generally include the end-points.

Let σ^*_{st} be the total number of *shortest feasible walks* (see Definition 1) from s to t, with $\sigma^*_{st}(v)$ be the total number of shortest feasible walks from s to t that pass through v. Then the (unnormalized) SOC-betweenness centrality of v, $BC^*(v)$, is given by

$$BC^*(v) := \sum_{s \neq v \neq t} \frac{\sigma^*_{st}(v)}{\sigma^*_{st}} \tag{6}$$

The computation of BC^* depends on counting the number of shortest feasible walks for each pair $s, t \in V$.

6.1 Counting shortest feasible walks

Let $d_G(u, v)$ for $u, v \in V$ be the geodesic distance from u to v. The term $\sigma_{st}(v)$ for $v \in V$ can be calculated as

$$\sigma_{st}(v) = \begin{cases} 0, & \text{if } d_G(s,t) < d_G(s,v) + d_G(v,t) \\ \sigma_{sv} \cdot \sigma_{vt}, & \text{otherwise} \end{cases} \tag{7}$$

This property, however, does not hold for counting shortest feasible walks in G. Thus, in order to count feasible walks in G, we turn to the directed graph $\mathscr{G}$.

Lemma 6.1. *Let $w = (s = u_0, u_1, \ldots, u_k = t)$ be an s-t walk in G of length k, with $u_i \in V$. w is a* feasible walk *in G if and only if there exists a walk in $\mathscr{G}$ with a node sequence of $(s, \kappa) = (u_0, i_0), (u_1, i_1), \ldots, (u_k, i_k) = (t, i_k)$ for some $0 \leq i_0, \ldots, i_k \leq \kappa$.*

Proof. For walk $w = (s = u_0, u_1, \ldots, u_k = t)$ in G, let $i_0, i_1, \ldots, i_k$ be the SOC value at nodes $u_0, \ldots, u_k$, respectively, during the walk. Since, w is a feasible walk, $i_0 = \kappa$ and node (u_{j+1}, i_{j+1}) is adjacent to (u_j, i_j) in $\mathscr{G}$, for $0 \leq j \leq k-1$. Therefore, the node sequence $(s, \kappa) = (u_0, i_0), (u_1, i_1), \ldots, (u_k, i_k) = (t, i_k)$ is a walk in $\mathscr{G}$. On the other hand, if $(s, \kappa) = (u_0, i_0), (u_1, i_1), \ldots, (u_k, i_k) = (t, i_k)$ is a walk in $\mathscr{G}$, then u_{j+1} is adjacent to u_j in G, for $0 \leq j \leq k-1$. Thus, $w = (s = u_0, u_1, \ldots, u_k = t)$ is a walk in G. □

For $0 \leq i \leq \kappa$, a walk from $(s, \kappa) \in \mathscr{V}$ to (t, i) can be viewed as a feasible walk from s to t in G. However, a shortest path from $(s, \kappa) \in \mathscr{V}$ to (t, i) is not necessarily a shortest feasible walk from s

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

to t in G. In order to count shortest feasible walks in G, we introduce a set of dummy nodes into $\mathscr{G}$ and call the new graph $\mathscr{G}_\star = (\mathscr{V}_\star, \mathscr{E}_\star)$, where

$$\mathscr{V}_\star := \mathscr{V} \cup \{ (u, \star) \mid u \in V\}$$
$$\mathscr{E}_\star := \mathscr{E} \cup \{ \big((u, i), (u, \star)\big) \mid (u, i) \in \mathscr{E}\}$$

Note: $\star$ can be viewed as a string or marker and is not a variable. The nodes (u, i) for $0 \le i \le \kappa$ represent node $u \in V$ at different states i. However, for a shortest feasible walk from s to t, we are interested in arriving at t at any state, thus introducing a dummy node $(t, \star)$ to capture all final states. The following lemma shows how adding these dummy nodes simplifies the process representing shortest feasible walks.

Lemma 6.2. *Let $w = (s = u_0, u_1, \ldots, u_k = t)$ be an s-t walk in G of length k, with $u_i \in V$. w is a* shortest *feasible walk in G if and only if there exists a* shortest path *in $\mathscr{G}_\star$ with a node sequence of $(s, \kappa) = (u_0, i_0), (u_1, i_1), \ldots, (u_k, i_k) = (t, i_k), (u_{k+1}, i_{k+1}) = (t, \star)$ for some $0 \le i_0, \ldots, i_k \le \kappa$; $k \in \mathbb{N}$.*

Proof. If w is a shortest feasible walk in G, from Lemma 6.1, it follows that there exists a walk w' in $\mathscr{G}$ and subsequently in $\mathscr{G}_\star$ with $w' = ((s, \kappa) = (u_0, i_0), (u_1, i_1), \ldots, (u_k, i_k) = (t, i_k))$. Suppose w' is not a path in $\mathscr{G}$. Then there exists a node (u_j, i_j) for some $j \in \mathbb{N}$ visited more than once. This forms a cycle C within w'. Define w'' as a node sequence in $\mathscr{G}$ where the cycle C in w' is replaced with (u_j, i_j). It is easy to see that w'' is a walk in $\mathscr{G}$ with length strictly less than w'. By Lemma 6.1, this implies that there exists a feasible walk in G with length less than w contradicting the assumption that w is a shortest feasible walk in G. Thus, w' and subsequently the node sequence $(s, \kappa) = (u_0, i_0), (u_1, i_1), \ldots, (u_k, i_k) = (t, i_k), (u_{k+1}, i_{k+1}) = (t, \star)$ for some $0 \le i_0, \ldots, i_k \le \kappa$, form a shortest path in $\mathscr{G}_\star$.

On the other hand, suppose w' is a shortest path from (s, κ) to $(t, \star)$ in $\mathscr{G}_\star$ of length $k + 1$ for some $k \in \mathbb{N}$. Lemma 6.1 implies that there exists an s-t feasible walk w in G of length k. By the same lemma, the existence of a shorter s-t feasible walk in G would imply the existence of a walk from (s, κ) to $(t, \star)$ in $\mathscr{G}_\star$ with length less than $k + 1$, contradicting the shortest path assumption. □

For a set $A \subset V$ and $s, t \in V$, define $\sigma_{st}(A)$ as the number of s-t shortest paths that pass through one or more nodes in A.

Lemma 6.3. *For $s, t, v \in V$, let $\gamma, \tau \in \mathscr{E}_\star$ with $\gamma = (s, \kappa)$, $\tau = (t, \star)$ and $A_v = \{(v, i) | 0 \le i \le \kappa\}$ then:*

1. $\sigma^*_{st} = \sigma_{\gamma\tau}$
2. $\sigma^*_{st}(v) = \sigma_{\gamma\tau}\big(A_v\big)$

Proof. For the first part, Lemma 6.2 gives a one-to-one mapping between set of shortest feasible walks in G and set of shortest paths in $\mathscr{G}_\star$ whose cardinalities are given by σ^*_{st} and $\sigma_{\gamma\tau}$, respectively. For the second part, $\sigma_{\gamma\tau}(A_v)$ counts the number of shortest paths in $\mathscr{G}_\star$ that pass through a node in A_v. Applying Lemma 6.2, any shortest path that passes through a node in A_v is a shortest feasible walk in G that passes through v. □

Due to the above lemma, we can compute $BC^*(v)$ as follows:

Theorem 1. *For graph $G = (V, E)$, and directed graph $\mathscr{G}_\star$, let $S = \{(s, i) \in \mathscr{E}_\star | s \in V, i = \kappa\}$ and $T = \{(t, i) \in \mathscr{E}_\star | s \in V, i = \star\}$, $A_v = \{(v, i) | 0 \le i \le \kappa\}$, with $v \in V$, then*

$$BC^*(v) = \sum_{\gamma,\tau \in \mathscr{E}_\star \gamma \in S, \tau \in T} \frac{\sigma_{\gamma\tau}\big(A_v\big)}{\sigma_{\gamma\tau}} \tag{8}$$

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

We now show how to compute $BC^*(v)$ without explicitly constructing $\mathscr{G}_\star$. In our computations, the value $\sigma_{\gamma\tau}(A_v)$ is approximated by $\sum_{i=0}^{\kappa} \sigma_{\gamma\tau}((v,i))$.

6.2 Computing SOC-betweenness centrality

In order to compute BC^*, we build on Brandes' algorithm (Brandes, 2001) for computing $BC(v)$. We first give a summary of Brandes' algorithm.

The *pair-dependency* is defined as the ratio

$$\delta_{st}(v) := \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{9}$$

of a pair $s, t \in V$ on an intermediary node $v \in V$. In order to eliminate the need for explicit summation of all pair-dependencies, Brandes introduces the notion of *dependency* of a vertex $s \in V$ on a single vertex $v \in V$, defined as

$$\delta_{s\bullet}(v) := \sum_{t \in V} \delta_{st}(v) \tag{10}$$

and shows that the dependency of $s \in V$ on any $v \in V$ obeys the following recursive relation:

$$\delta_{s\bullet}(v) = \sum_{w: v \in P_s(w)} \frac{\sigma_{sv}}{\sigma_{sw}} \cdot (1 + \delta_{s\bullet}(w)) \tag{11}$$

where $P_s(v)$ is the set of *predecessors* of a vertex v during a breadth-first search (BFS) from source $s \in V$. It is given by

$$P_s(v) := \{u \in V : \{u, v\} \in E,\ d_G(s, v) = d_G(s, u) + 1\} \tag{12}$$

where $d_G(s, v)$ is the geodesic distance from s to v. In summary, Brandes' algorithm for computing BC is as follows: for each source node, $s \in V$,

i. perform BFS computing number of shortest paths to every other node, $t \in V$
ii. back propagation: compute $\delta_{s\bullet}(v)$ for $v \in V$ in order of non-increasing distance from s.

One major difference between equation (8) and the standard computation of BC is that Equation (8) is constrained by the fact that the source and target nodes must be chosen from sets S and T. Therefore, the recursive relation given by Equation (11) cannot be used as it is because not all nodes are target nodes. For $T \subset V$, consider the function

$$\delta^T_{s\bullet}(v) := \sum_{t \in T} \delta_{st}(v) \tag{13}$$

then

Lemma 6.4.

$$\delta^T_{s\bullet}(v) = \sum_{w: v \in P_s(w)} \frac{\sigma_{sv}}{\sigma_{sw}} \cdot (\mathbb{1}_T(w) + \delta^T_{s\bullet}(w)) \tag{14}$$

where $\mathbb{1}_T(w)$ is the indicator function such that $\mathbb{1}_T(w) = 1$ if $w \in T$ and 0 otherwise.

Proof. For each term on the right side, if $w \in T$, then the summand follows from Equation (11). Consider the case if $w \notin T$. Extend the definition of the pair-dependency to include an edge e such that, $\delta_{st}(v, e) := \sigma_{st}(v, e)/\sigma_{st}$, where $\sigma_{st}(v, e)$ is the number of shortest s-t paths that contain both v and e. Then Brandes showed that

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

$$\delta_{s\bullet}(v) = \sum_{w:v\in P_s(w)} \sum_{t\in V} \delta_{st}(v, \{v, w\})$$

and

$$\delta_{st}(v, \{v, w\}) = \begin{cases} \frac{\sigma_{sv}}{\sigma_{sw}}, & \text{if } t = w \\ \frac{\sigma_{sv}}{\sigma_{sw}} \cdot \frac{\sigma_{st}(v)}{\sigma_{st}}, & \text{otherwise} \end{cases}$$

It then follows that

$$\delta^T_{s\bullet}(v) = \sum_{w:v\in P_s(w)} \sum_{t\in T} \delta_{st}(v, \{v, w\})$$

So for $w \notin T$, then $t \neq w$ and

$$\begin{aligned} \sum_{w:v\in P_s(w)} \sum_{t\in T} \delta_{st}(v, \{v, w\}) &= \sum_{w:v\in P_s(w)} \frac{\sigma_{sv}}{\sigma_{sw}} \cdot \frac{\sigma_{st}(w)}{\sigma_{st}} \\ &= \sum_{w:v\in P_s(w)} \frac{\sigma_{sv}}{\sigma_{sw}} \cdot \delta s\bullet^T(w) \end{aligned}$$

□

The recursive relation in Equation (14) is used to compute the SOC-betweenness centrality. Algorithm (1) describes this computation in detail.

7. Random-walk betweenness centrality

A common criticism for betweenness centrality is that it does not take non-shortest paths into account and is therefore inappropriate in cases where information spread is governed by other rules (Borgatti, 2005). As a result, variants of betweenness centrality have been proposed such as betweenness measures based on network flow (Freeman et al., 1991) and RWBC (Newman, 2005; Brandes & Fleischer, 2005). In some sense, as suggested by Newman, "RWBC and BC can be viewed as being on opposite ends of a spectrum of possibilities, one representing information that is moving at random and has no idea of where it is going and the other knowing precisely where it is going." Some real-world situations mimic these extremes (Newman, 2005; Freeman, 1978), however, others, such as the small-world experiment (Kleinfeld, 2002), fall somewhere in between.

In a network where the flow process is coupled with an SOC constraint, it is therefore natural to also propose a variant of RWBC for such networks. If we consider an undirected connected graph, for any pair of nodes s, t, a random walk starting at s will eventually arrive at t with high probability. However, in a network where the flow is coupled with a SOC constraint, and likewise, a directed network that is not strongly connected, not every random walk starting at s has a positive probability of arriving at t. With this in mind, the proposed variant of RWBC only considers walks that arrive at the destination node. For example, if a node does not have enough SOC to travel from s to t via any walk, then the pair s-t does not contribute to centrality score.

Consider RWBC proposed in Newman (2005). Unlike the standard betweenness centrality measure that only considers shortest paths between a pair of nodes, RWBC takes all paths into account while giving more importance to shorter paths. RWBC of a node i is defined as the *net* number of times a random walk passes through i. By *net*, authors meant that if a walk passes through i and later passes back through it in the opposite direction, the two would cancel out and there is no contribution to the betweenness.

RWBC was originally proposed for undirected graphs. In this section, we first generalize RWBC to directed graphs. In a directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, for any pair of nodes $s, t \in \mathscr{V}$, it is not guaranteed that every random walk from s will eventually arrive at t. We generalize RWBC

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

Algorithm 1 SOC-Betweenness Centrality

Input: $G = (V, E)$, $\mathcal{V}_\star$, Ω, κ
Output: $bc[v]$, $v \in \mathcal{V}_\star$
$bc[v] \leftarrow 0$, $v \in \mathcal{V}_\star$
$\Sigma \leftarrow \{(u, i) \in \mathcal{V}_\star | u \in \mathcal{V}_\star, i = \kappa\}$
for $s \in \Sigma$ **do**
    $S \leftarrow$ empty stack;
    $P[\omega] \leftarrow$ empty list, $\omega \in \mathcal{V}_\star$;
    $\sigma[t] \leftarrow 0$, $t \in \mathcal{V}_\star$;$\sigma[s] \leftarrow 1$
    $d[t] \leftarrow -1$, $t \in \mathcal{V}_\star$; $d[s] \leftarrow 0$
    $Q \leftarrow$ empty queue;
    enqueue $s \rightarrow Q$;
    while Q *not empty* **do**
        dequeue $(v, i) \leftarrow Q$;
        push $(v, i) \rightarrow S$;
        if $i \neq \star$ **then**
            for *neighbor* w *of* v **do**
                if $w \in \Omega$ **then** $current_node \leftarrow (w, \kappa)$
                else
                    if $i \geq 0$ **then** $current_node \leftarrow (w, i - 1)$
                    else $current_node \leftarrow -1$
                if $current_node \neq -1$ **then**
                    if $d[current_node] < 0$ **then** ▷ *w found for first time*?
                        enqueue $current_node \rightarrow Q$;
                        $d[current_node] \leftarrow d[(v, i)] + 1$;
                    if $d[current_node] = d[(v, i)] + 1$ **then** ▷ shortest path to w via v?
                        $\sigma[current_node] \leftarrow \sigma[current_node] + \sigma[(v, i)]$;
                        append $(v, i) \rightarrow P[current_node]$;
    $\delta[v] \leftarrow 0$, $v \in \mathcal{V}_\star$; ▷ S return vertices in order of non-increasing distance from s
    $\chi[(v, i)] \leftarrow 0$, $(v, i) \in \mathcal{V}_\star$
    $\chi[(v, i)] \leftarrow 1$, $(v, \star) \in \mathcal{V}_\star$
    while S *not empty* **do**
        pop $(w, i) \leftarrow S$;
        if $\chi[(w, i)] = 1$ **then**
            for $(v, j) \in P[(w, i)]$ **do** $\chi[(v, j)] \leftarrow 1$
                if $i = \star$ **then** $\delta[(v, j)] \leftarrow \delta[(v, j)] + \frac{\sigma[(v,j)]}{\sigma[(w,i)]} \cdot (1 + \delta[(w, i)])$;
                else $\delta[(v, j)] \leftarrow \delta[(v, j)] + \frac{\sigma[(v,j)]}{\sigma[(w,i)]} \cdot \delta[(w, i)]$;
            if $(w, i) \neq s$ **then** $bc[(w, i)] \leftarrow bc[(w, i)] + \delta[(w, i)]$

for directed graphs to only include random walks from s to t. Let $\mathcal{G}_{s,t} = (\mathcal{V}_{s,t}, \mathcal{E}_{s,t})$ be a subgraph of $\mathcal{G}$ such that every node lies on a walk from s to t. RWBC is adjusted for $\vec{\mathcal{G}}_{s,t}$ as follows. Let $\mathcal{A}$ adjacency matrix with D the out-degree diagonal matrix, where D is defined as

$$[D]_{ij} := \begin{cases} deg_+(v_i) & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

where $deg_+(v_i)$ is the out-degree of node v_i. Define the transition matrix of $\vec{\mathcal{G}}_{s,t}$ as

$$M := D^{-1}\mathcal{A} \tag{15}$$

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

For a walk starting at s, the probability that it is at j after r steps is given by $[M^r]_{sj}$. The probability that the walk continues further to an adjacent vertex i is $[M^r]_{sj}d_j^{-1}$, where d_j^{-1} is the out-degree at j. Thus, the expected number of times a walk from s to t uses the directed edge (j, i) is given by $[(I - M_t)^{-1}]_{sj}d_j^{-1}$, which is the s-jth entry of the matrix given by

$$(I - M_t)^{-1}D_t^{-1} = (D_t - A_t)^{-1} \tag{16}$$

where D_t and A_t are the matrices derived from deleting row and column t. Add the zero column back to $(D_t - A_t)^{-1}$ and call this matrix T. Let $\mathbf{s}$ be the vector given by

$$s_i := \begin{cases} 1, & \text{if } i = s \\ -1, & \text{if } i = t \\ 0, & \text{otherwise} \end{cases}$$

Let the vector f be defined as

$$f := \mathbf{s}^T T$$

then, the ith entry of f, f_i, represents the expected number of walks from s to t that pass through node i. If $\mathcal{D}_f$ is the diagonal matrix with f_i at the ith diagonal position, then the matrix

$$\mathcal{F} := \mathcal{D}_f\mathcal{A}$$

gives a matrix whose i-j value represents the expected number of times a random walk from s to t uses edge (i, j). The net flow of random walk through the ith vertex is for a given s-t pair is given by

$$I_i^{(st)} = \frac{1}{2}\sum_{(i,j)\in\mathscr{E}_{s,t}} \left|\mathcal{F}_{i,j} - \mathcal{F}_{j,i}\right| \tag{17}$$

The expression in (17) is used to compute the centrality scores of a directed graph $\mathscr{G}_{s,t}$ arising from random walks starting at node s to t. Note that for every node v in $\mathscr{G}_{s,t}$, v must be at a finite distance from s and t. The centrality for each node in $\mathscr{G}$ is then given by the sum of the individual scores for each source–target pair. Let $\hat{Y}$ be the vector of centrality scores of $\mathscr{G}$; the SOC-RWBC is given by the vector

$$\hat{Y}^T \cdot \mathcal{I}_\kappa \tag{18}$$

where $\mathcal{I}_\kappa$ is the block matrix defined in (2).

8. Computational experiments

In the preceding sections, mathematical models for the three proposed centrality measures are given. The question then arises, "How good are these centrality measures?" We tackle this question from three different perspectives. First, *usability*: how can we meaningfully use the proposed centrality measures. Second, *robustness*: how robust are the centrality measures with respect to their parameters. Lastly, *novelty*: how are the proposed centrality measures different from their well-established predecessors. Experiments in this section are carried on the graph data sets given in Table 1.

8.1 Usability

In Borgatti (2005), the expected centrality is defined as a centrality score given by a closed-form expression, and *realized centrality* as the actual centrality score observed in the context of a particular flow process. Therefore, one can view a centrality measure as a formula-based prediction of

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

Table 1. Experimental graph data sets: $d_{\min}, d_{avg}, d_{\max}$ represents the minimum, average, and maximum degree.

Graph	Nodes	Edges	$d_{\min}$	d_{avg}	$d_{\max}$	Reference
Router Network	2,114	6,632	1	6	109	(Rossi & Ahmed, 2015)
Minnesota State Road Network	2,642	3,303	1	2	5	(Davis & Hu, 2011)
Gnutella Network	6,301	2,077	1	7	97	(Ripeanu & Foster, 2002)
Collaboration Network	5,242	14,496	0	5	81	(Leskovec et al., 2007b)

a flow process through a node. It is therefore important to compare the predictions given by the closed-form expression with the actual frequency of traffic observed flowing through a node across multiple instances. For example, in order to test whether betweenness centrality is a good prediction of observed traffic through a node, the expected betweenness centrality is compared with the realized betweenness centrality, where the actual frequency of traffic through a given node is referred to as the realized betweenness centrality, while the formula-based centrality is referred to as the expected betweenness centrality. In this section, we compare the expected centrality with the realized centrality values for the proposed centrality measures. Given that realized centrality scores are achieved by running long simulations, we show the usability of the proposed measures as way to efficiently estimate the outcome of these computationally expensive simulations.

In order to observe realized centrality values, simulations for each of the three flow process are developed. In general, centrality measures are primarily used either as ranking algorithms or as methods for identification of influential nodes. We therefore compare the expected and realized centrality values using Kendall's Tau (Kendall, 1938) rank correlation coefficient. Kendall's Tau is given by

$$\tau := \frac{2}{n(n-1)} \sum_{i<j} sgn\Big[(y_i - y_j)(z_i - z_j)\Big]$$

where, for each node i, we denote the node's spreading influence and its centrality measure by y_i and z_i, respectively. The $sgn(y)$ is a piecewise function such that $sgn(y) = 1$ if $y > 0$, -1 if $y < 0$ and 0 if $y = 0$. The values of τ belong to the range $[-1, 1]$, where larger values of τ correspond to a higher correlation between the expected and realized centralities.

8.1.1 SOC Katz centrality

In order to compute the realized centrality values with respect to SOC-Katz centrality, we turn to the susceptible–infected–recovered (SIR) spreading model (also called susceptible–infected–removed model) (Hethcote, 2000). Klemm et al. (2012) suggested that the eigenvector centrality can be used for estimating a spreading influence of the nodes in the SIR model, by (Liu et al., 2016) defining the dynamical-sensitive (DS) centrality and showing that it more accurately locates influential nodes in the SIR model. The DS centrality is very closely related to the Katz centrality.

In the SIR model, a node can be in one of following states: (i) susceptible, nodes can become infected, (ii) infected, nodes are infected and can infect susceptible nodes, and (iii) recovered, nodes have recovered and developed immunity, thus cannot be infected again. In order to estimate the spreading influence of a node v, initially, all nodes are susceptible and v is infected. At each step, an infected node tries to infect its susceptible neighbors and succeeds with probability α. The infected node enters the recovered state with probability μ. In this work, we set $\mu = 1$, that is, the infection can be transmitted only once. The process stops if no new infections are formed or after a fixed number of steps. We generalize the SIR spreading process to accommodate the SOC parameter.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

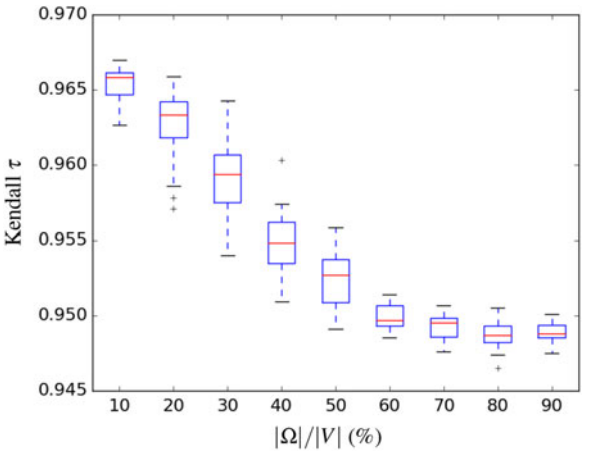


Figure 1. Comparison of nodes' spreading influence according to the generalized SIR model and SOC-Katz centrality on the routers network. Each boxplot represents 30 random choices of the set Ω with spreading probability $\alpha = 0.03$, and $\kappa = 5$.

Define an edge (i, j) as *active* if node i infected j via edge (i, j). A stopping criteria for the SIR spreading process for a fixed number of steps κ can be viewed as follows: Let u be the initially infected node, then

> *an infected node v cannot infect its susceptible neighbors if there exists a path of length κ from u to v consisting of only active edges*, that is, the infection dies out after κ steps.

In order to generalize the SIR model to accommodate a flow process based on SOC, we modify the above stopping criteria to:

> *an infected node v cannot infect its susceptible neighbors if there exists a path of length κ from either u, or non-susceptible $w \in \Omega$ to v consisting of only active edges.*

If the set $\Omega \subset V$ is empty, then SOC-Katz centrality is equivalent to the DS centrality which is shown in Liu et al. (2016) to be highly correlated with the nodes' spreading influence according to the SIR model.

Experiments are carried out to show that the above generalized SIR spreading process is highly correlated with the proposed SOC-Katz centrality. For this experiment, we use the network representing the Internet at the major router level (Rossi & Ahmed, 2015; Spring et al., 2002) consisting of 2,114 nodes and 6,632 edges. The nodes and edges represent routers and the connections between them, respectively. We set $\kappa = 5$, and $\alpha = 0.03$, and vary the size of the set $\Omega \subset V$ such that the ratio $|\Omega|/|V|$ ranges from 0.1 to 0.9. For each value of $|\Omega|/|V|$, the set Ω is chosen at random, and the corresponding spreading influence is estimated for each node by running the generalized SIR model 10^4 times. This is repeated 30 times. The boxplot in Figure 1 shows the correlation between the spreading influence as a result of the generalized SIR model compared to SOC-Katz centrality. The results show Kendall Tau correlation values in the range $(0.945, 0.970)$ suggesting that the two processes are very highly correlated.

8.1.2 SOC-(random-walk) betweenness centrality

To demonstrate the SOC-RWBC and SOC-betweenness centralities, we experiment with two networks, namely, a computer network and road network. The computer network is generated from

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

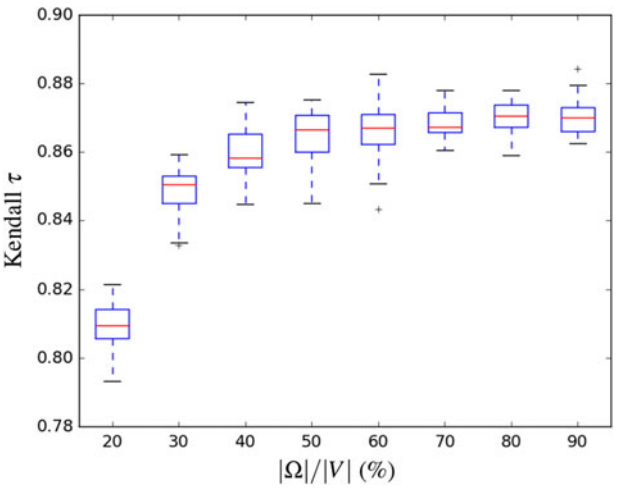


Figure 2. Correlation scores for expected versus realized centrality for SOC-betweenness centrality for the Gnutella network. Each boxplot represents 30 random choices of the set Ω, with $\kappa = 4$.

the P2P network, Gnutella (Leskovec et al., 2007b; Ripeanu & Foster, 2002), and consists of 6,301 nodes and 20,777 edges. The road network (Davis & Hu, 2011) represents Minnesota state roads and consists of 2,642 nodes and 3,303 edges. We simulate traffic on both networks.

The realized centralities are computed using the particle hopping. The particle hopping model is a method used in vehicular flow theory (Nagel, 1996). In this model, a section of a road is represented by a node and a vehicle as a particle where each node can only be occupied by one particle at a given time. This model is sometimes referred to as cellular automata and gives a minimal model for traffic flow behaviors (Huitema, 2000). The flow of packets through the internet has also been modeled by cellular automata (Liu et al., 2002; Huisinga et al., 2001). In Holme (2003), the fraction of time steps that node is occupied by a particle is referred to as the *occupation ratio*.

For the application to electric vehicles, the value κ represents the number of steps the car can travel before its battery runs out of charge. For a message or vehicle being propagated from node s to t, we simulate the traffic on the nodes when a routing algorithm propagates the message or vehicle in one of the two cases: (i) via a shortest feasible walk and (ii) via a random feasible walk. We add the condition that the routing algorithm is informed and takes the current κ counter of the message or vehicle, and target t, into account before deciding which neighbor to direct it to. In other words, if a message or vehicle cannot be successfully propagated to its destination due to the value of κ, then the message or vehicle is not propagated at all and therefore does not contribute to the traffic of the network.

Experiments are carried on the Gnutella network, where we set $TTL = 4$ and record the occupation ratio based on the corresponding routing algorithms. The occupation ratio is then compared to the proposed centrality measures. The results for SOC-betweenness centrality and SOC-RWBC are presented in Figures 2 and 3, respectively. The results show Kendall Tau values in the range (0.79, 0.82) for a ratio $|\Omega|/|V|$ of 0.2 and (0.86, 0.88) for a ratio $|\Omega|/|V|$ of 0.9 for SOC-betweenness centrality. Similar correlation scores and trends for SOC-RWBC are observed suggesting a high correlation between the expected centralities and realized centrality measures.

For experiments on the Minnesota state road network, we set $\kappa = 20$. We choose a relatively larger value of κ for the road network experiments because we assume that electric vehicles can travel a relatively long distance if it starts fully charged. As in the Gnutella experiments, we record the occupation ratio. The results for SOC-betweenness centrality are presented in Figure 4. The results show Kendall Tau values in the range (0.79, 0.86) for a ratio $|\Omega|/|V| = 0.2$ and (0.83, 0.87) for a ratio $|\Omega|/|V| = 0.9$ for SOC-betweenness centrality.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

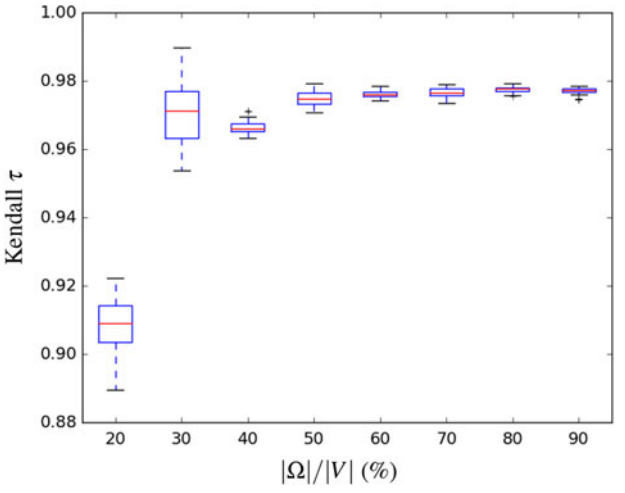


Figure 3. Correlation scores for expected versus realized centrality for SOC RWBC for the Gnutella network. Each boxplot represents 30 random choices of the set Ω, with $\kappa = 4$.

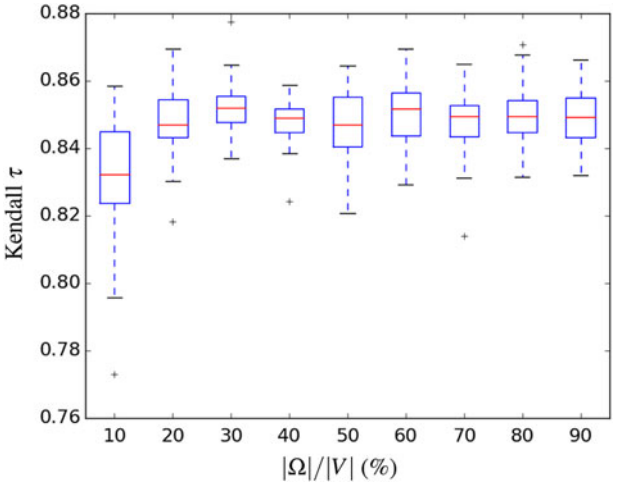


Figure 4. Correlation scores for expected versus realized centrality for SOC-betweenness centrality for the Minnesota state road network. Each boxplot represents 30 random choices of the set Ω, with $\kappa = 20$.

8.2 Robustness and novelty

The parameter κ is application dependent, so it is important to understand how the proposed centrality measures behave for different values of κ. From the mathematical expressions of our novel centrality measures, it is clear that for a large enough κ, the proposed centrality measures would become identical to their baseline centrality measures as in this case, the limitation of SOC-dependent distance is gradually vanishing. In this section, we carry out experiments to understand how the proposed measures compare to their baseline measures while varying the parameter κ. The goal of the experiments is to quantify what is not captured when using the well-established centrality measures for given values of κ, thus demonstrating the robustness of the results and novelty of the measures.

The first set of experiments is carried out on toy graphs to illustrate the difference in central nodes when using the proposed centrality measures versus their baseline measures. The first toy graph is a graph formed by connecting two 5×5 grid graphs by a path of length 5. The second is a 10×10 grid graph. The second set of experiments uses real-world data sets. In these experiments,

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

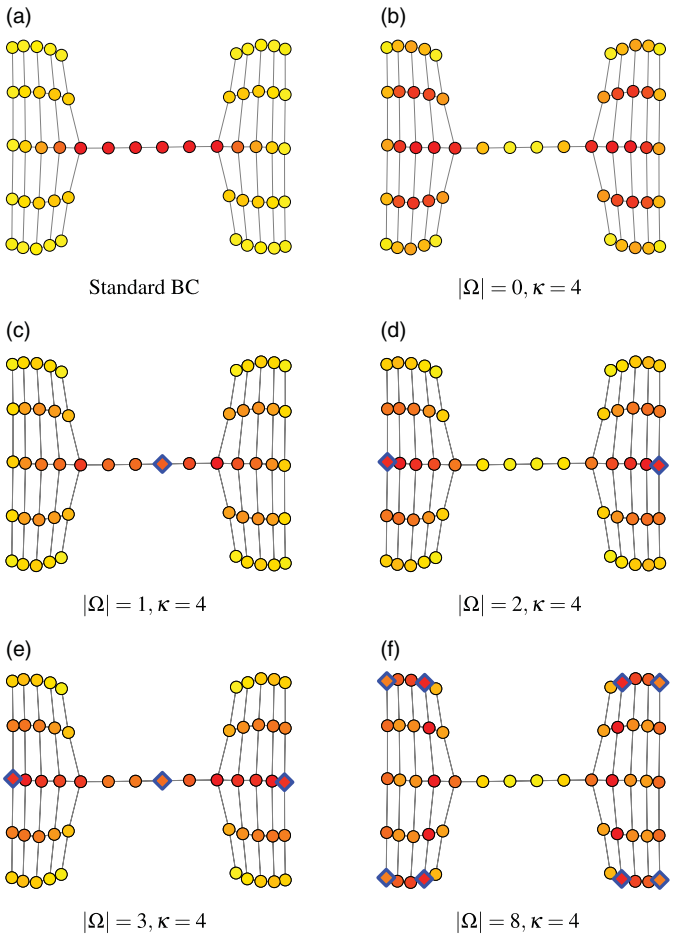


Figure 5. Comparison of BC and SOC-BC. The blue diamond-shaped nodes represent nodes in Ω. In (a), the nodes connecting the two components of the graph, which can be viewed as a "bridge" between two communities, have high centrality scores because they are essential for flow from one component to another. Whereas in (b)–(f), since the flow is limited to just four steps, the majoring of the flow in the graph would be within the two components thus the bridge nodes are no longer as important.

we focus on the Minnesota state road network and a collaboration network constructed using the scientific collaboration data (Leskovec et al., 2007b), consisting of 5,242 nodes and 14,496 edges.

The experiments on toy graphs are used to visually illustrate to the reader the difference between the proposed centrality measures and their baseline measures, thus providing an intuition on how the measures work. The difference between SOC-BC and BC for small values of κ is illustrated in Figure 5. In this experiment, we set $\kappa = 4$. We use a color spectrum from red to yellow, showing the most central to the least central nodes respectively. The graph in Figure 5(a) represents the standard BC. As expected, the nodes along the bridge are the most central nodes. The graphs in Figure 5(b)–(f) show different scenarios where the nodes in Ω are marked with a blue-edge diamond-shaped node. As we can see in Figure 5(d) and (f), depending on the value of κ and nodes in Ω, the centrality scores can be significantly different from the standard BC, where the most central nodes based on BC are now among the least central nodes based on SOC-BC.

Differences between SOC-Katz and Katz centrality are illustrated in Figure 6. In this experiment, we set $\kappa = 4$. The graph in Figure 6(a) represents the standard Katz centrality. As expected, the nodes toward the center of the grid are the most important nodes according to this model.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

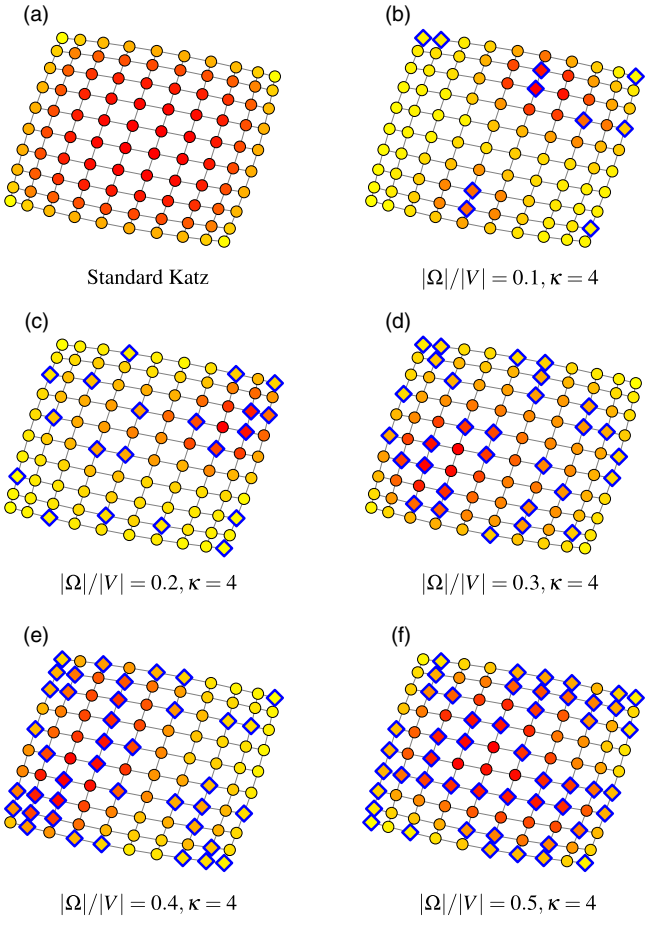


Figure 6. Comparison of Katz and SOC-Katz; The blue diamond-shaped nodes represent nodes in Ω. In (a), the standard Katz centrality shows that the nodes more centrally located on the grid have a higher importance. For example, with respect to information spreading in a network, this implies that these nodes are the most influential in information spreading. However as observed from (b) to (f) if the information spread has fixed travel distance, then just the connectivity structure of the network is not enough to conclude about the most influential nodes.

The graphs in Figure 6(b)–(f) show different scenarios where the nodes in Ω are marked with a blue-edge diamond-shaped node. As we can see in Figure 6(e), depending on the value of κ, even with a relatively large ration of $|\Omega|/|V|$ (0.5 for (e)), the centrality scores can still be significantly different than the scores from their baseline models.

A more comprehensive study comparing SOC-Katz with Katz on the grid graph is shown in Figure 7. For each value of κ, with $2 \leq \kappa \leq 16$, we run 30 experiments. Each experiment consists of choosing nodes at random to be in the set Ω. The boxplots in blue represent experiments with $|\Omega|/|V| = 0.1$, while $|\Omega|/|V| = 0.2$ are represented in red. We choose the values 0.1 and 0.2 for the ratio $|\Omega|/|V|$ because in most applications the set Ω will be considerably smaller than V. For example in social networks, the percentage of lurkers in an online community is estimated to range from 50% to 90% of the total membership (Katz, 1998; Mason, 1999; Soroka et al., 2003). As expected, the results show that as κ increases, the correlation between SOC-Katz and Katz ranking increases. It is interesting to observe that since Katz centrality is based on infinite-length random walks emanating from a node, it is not clear what value of κ would make SOC-Katz identical to the standard Katz centrality for a given graph. However, this is not the case with betweenness centrality which is based on shortest paths. For a given graph, setting κ to the longest shortest path would make SOC-BC identical to BC.

The second set of experiments in this section is carried out on real-world graphs, the Minnesota state road network, and the collaboration network. In the Minnesota state road network, first, we compare SOC-BC with BC, while varying the value of κ; second, we compare SOC-RWBC and

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

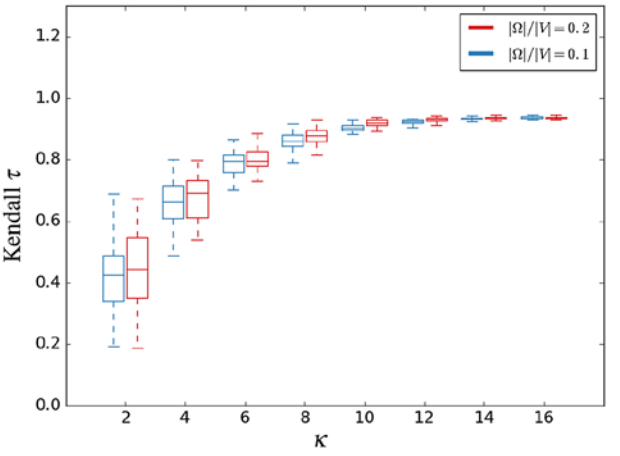


Figure 7. SOC-Katz versus Katz for 10×10 grid graph.

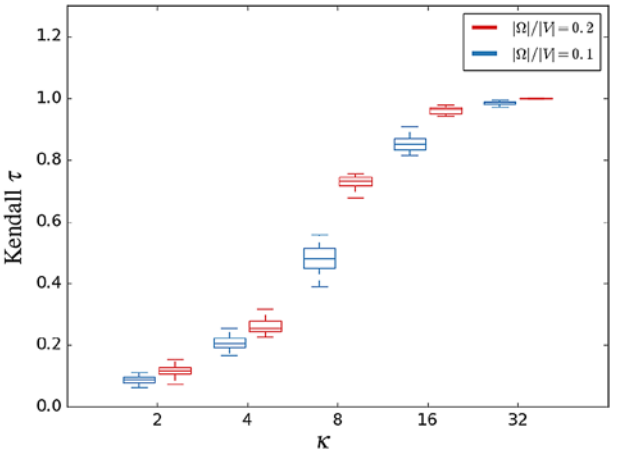


Figure 8. SOC-BC versus BC for Minnesota state road network.

RWBC for a given source–target pair. The comparison of SOC-BC and BC is shown in Figure 8. The parameter κ is varied from 2 to 16. For each value of κ, we perform 30 experiments where each experiment consists of sampling a set of nodes $\Omega \subset V$ for a fixed ratio $|\Omega|/|V|$. We fix the ratio to 0.1 and 0.2, represented by blue and red boxplots, respectively. Given that the Minnesota state road network has an average shortest path length of approximately 35.4, the results show that for values of κ, smaller than the average shortest path length, we can get significant differences between SOC-BC and BC ranking. Thus, we find where the standard BC may potentially fail in identifying central nodes. We visually demonstrate a similar result for SOC-RWBC and RWBC in Figure 9. In this experiment, we pick a pair of nodes representing a source and target and then compute the RWBC scores contributed by the two nodes referring to them as *s*-*t*-RWBC. Given that RWBC is identical to the current flow betweenness centrality (Brandes & Fleischer, 2005), one can think of this experiment as injecting a unit of current from the source flowing to target and measuring the fraction of current flowing through each node. The graph in Figure 9(a) represents the *s*-*t*-RWBC scores for the source–target pair represented by nodes in black. The *s*-*t*-RWBC values identical to zero are represented with nodes with negligible size. As expected, the results show higher *s*-*t*-RWBC values for nodes close to the source and target. We perform a similar experiment for SOC-RWBC with $\kappa = 20$, while the distance from source to target is larger than 20. This implies that every random walk from source to target must pass through at least one node in Ω. The results in Figure 9(b) show the *s*-*t*-SOC-RWBC values for the given source–target pair. In the case of SOC-RWBC, the higher central nodes are now the nodes close to the nodes in Ω.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

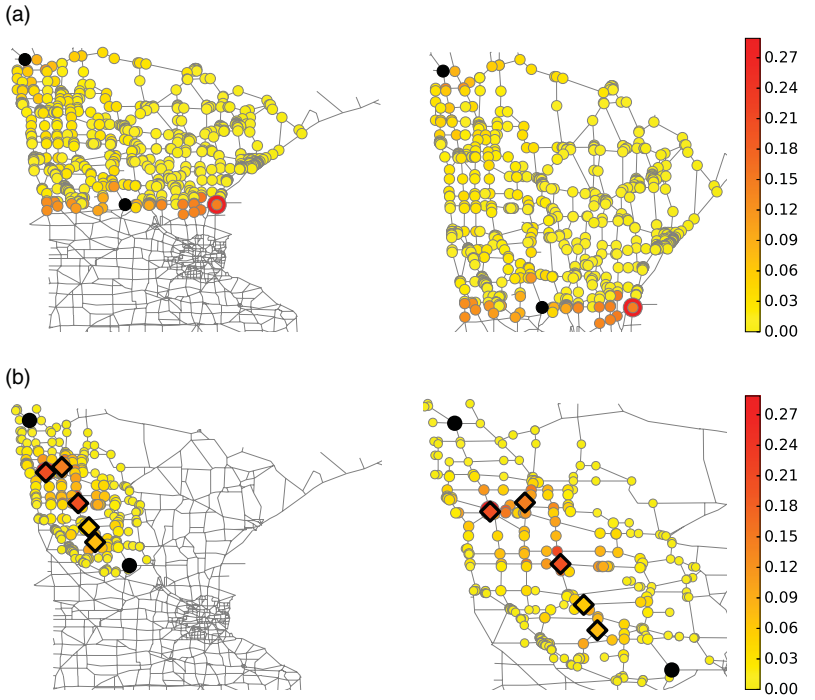


Figure 9. Comparison of RWBC with SOC-RWBC for a single *s*-*t* pair over the Minnesota state road network. The top-left and center black nodes represent the source and target nodes, respectively. Nodes with *s*-*t*-centrality scores equal to 0 have negligible node sizes. (a) RWBC for a given *s*-*t* pair. (b) SOC-RWBC for a given *s*-*t* pair. Nodes with a triangular marker represent nodes in Ω.

This example demonstrates how SOC-RWBC can be used to identify congested nodes in a road network that is equipped with wireless charging lanes.

A comparison of SOC-Katz versus standard Katz with variations of the parameter κ on the Collaboration network is shown in Figure 10. In general, for different values of κ, the correlation of SOC-Katz and Katz is high, generally above 0.8. However, once we plot the different ranking we see significant differences with the node rankings of the two measures. Thus, Kendall Tau correlation does not give a complete picture for this network. In particular, a node that is ranked highly with the standard Katz centrality can have a significantly less rank in a ranking with SOC-Katz. However, conversely, highly ranked nodes with SOC-Katz generally tend to also be highly ranked with respect to the standard Katz. With respect to the application to posters and lurkers in social networks, this follows the intuition that a user with a large number of neighbors (friends) can still be non-influential if the user together with all his/her neighbors (friends) are lurkers. On the other hand, a highly influential node would generally have many neighbors (friends).

9. Conclusion

An estimation of node spreading influence in a network is an important step toward understanding and controlling the spreading dynamics over the network. Centrality measures are traditionally used to identify influential nodes in a network. In this work, we extend the well-known measures of Katz, betweenness, and RWBC to models that accommodate a resource, necessary for the spread, being consumed along the way. We present algorithms to compute the proposed centrality measures and carry out experiments on real-world networks. Lastly, we demonstrate simulation models that describe the flow process and show that they are highly

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

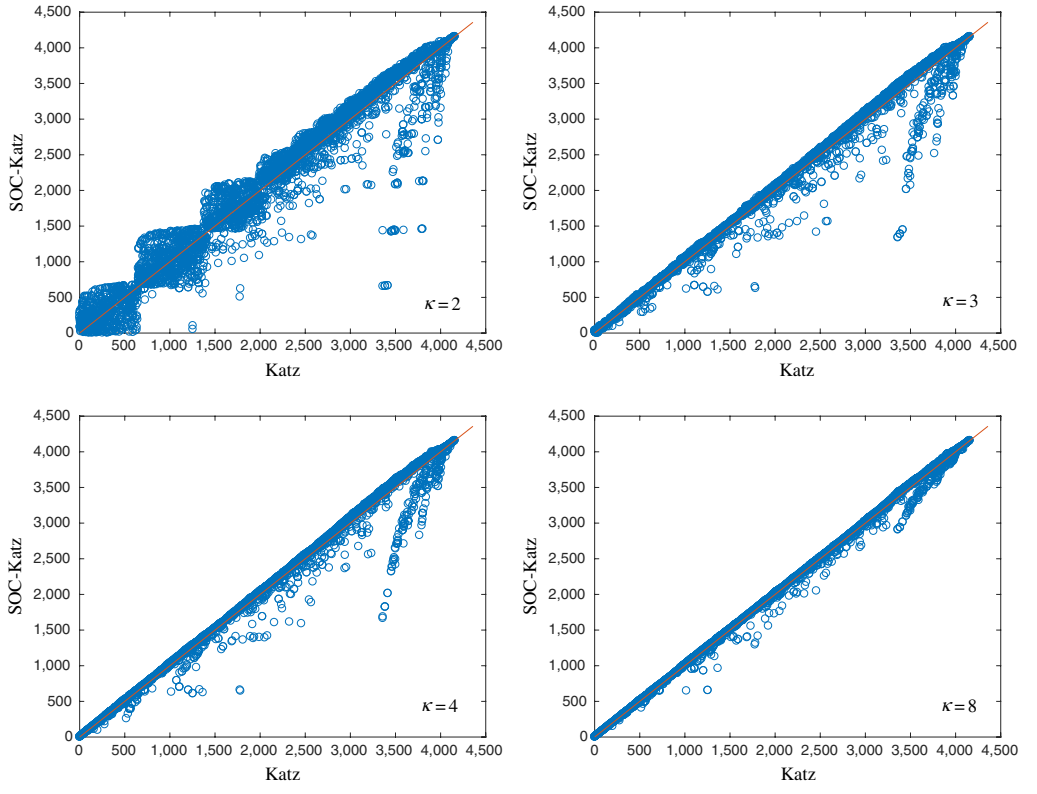


Figure 10. Comparison of node rankings based on SOC-Katz and Katz for different values of κ with $|\Omega|/|V| = 0.1$. Even if the above ranking gives high Kendall Tau correlation, we notice that with the introduction of a ranking-based SOC-Katz, highly ranked Katz nodes can significantly lose their ranking; however, less important nodes do not significantly increase with rank.

correlated with the proposed centralities. In order to answer the question, "How good are these centrality measures?", we analyze the centrality measures from three different perspectives, namely, usability, robustness, and novelty. From the usability perspective, among other experiments, we demonstrate how the centrality measures can be used to identify congested nodes in computer and road networks. From the robustness and novelty perspective, in the application of posters and lurkers in a social network, we showed that the proposed extension of Katz centrality follows the intuition that a user with a large number of neighbors (friends) can still be non-influential if the user together with all his/her neighbors (connections) are lurkers. On the other hand, a highly influential node would generally have many neighbors (connections).

The proposed measures take into account a spreading process that depends on a resource, such that the spreading process would be impossible without. Our numerical experiments demonstrate that the proposed measures differ significantly from the original measures when the resource is limited. On the other hand, they become identical to the original measures as the quantity of resource available tends to infinity. As a result, the proposed measures give a new tool and perspective to different application domains. For example, in the application of a road network equipped with wireless charging lanes, an optimal placement of these lanes with respect to traffic distribution could be one where the distribution of centrality scores of all nodes is taken into account. In another domain, high centrality nodes can be considered for targeted attack or immunization strategies.

For a given application, the choice of which centrality measure to use to draw a conclusion about the network is extremely important as using a wrong measure can lead to meaningless results. The measures of SOC-RWBC and SOC-Katz are both based on random walks on

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

the network. It is, however, important to note that, just as the standard measures, the random walks associated with SOC-RWBC have a fixed source and target node, while the random walks associated with SOC-Katz only have a fixed source node. Thus SOC-Katz is more suitable for applications where the flow process does not have a specified destination, for example, a disease spread. The SOC-RWBC and subsequently SOC-BC are suitable for applications where the flow process has a specified destination, for example, vehicle flow.

There are numerous future research directions associated with the resource consumption-based centralities. For example, we propose to explore other fundamental network properties such as connectedness, clustering, and network robustness in the context of consumable resource networks. In this work, there exists a set of nodes that facilitate flow in the network; conversely, problems in network interdiction (Wood, 1993) deal with the identification of nodes that hinder flow. An interesting direction is to explore the relationship between these two problems in more detail. Another highly relevant direction for the future work is to consider a distribution of resource consumption-based centralities in realistic network generation (Gutfraind et al., 2015; Staudt et al., 2017) . This is particularly important for the simulation and verification studies. To the best of our knowledge, no generating model currently considers a distribution of resource consumption-based centralities. Also, the resource consumption models can be generalized for clusters and communities. Moreover, one of the natural extensions of this work is introducing resource consumption element to node and edge similarity measures.

Acknowledgments. We thank the anonymous reviewers whose comments and suggestions helped to improve and clarify this manuscript. This research is supported by the National Science Foundation under Award 1647361. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Conflict of interest. The authors have nothing to disclose.

References

Albert, R., & Barabási, A-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), 47.

Albert, R., Jeong, H., & Barabási, A-L. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378.

Altshuler, Y., Puzis, R., Elovici, Y., Bekhor, S., & Pentland, A. S. (2011). *Augmented betweenness centrality for mobility prediction in transportation networks*. Paper presented at the International Workshop on Finding Patterns of Human Behaviors in Networks and Mobility Data, nemo11.

Bae, J., & Kim, S. (2014). Identifying and ranking influential spreaders in complex networks by neighborhood coreness. *Physica A: Statistical Mechanics and Its Applications*, *395*, 549–559.

Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks*. Cambridge University Press.

Basu, A., Fleming, S., Stanier, J., Naicken, S., Wakeman, I., & Gurbani, V. K. (2013). The state of peer-to-peer network simulators. *ACM Computing Surveys (CSUR)*, *45*(4), 46.

Bavelas, A. (1948). A mathematical model for group structures. *Human Organization*, *7*(3), 16–30.

Bettencourt, L. M. A., Cintrón-Arias, A., Kaiser, D. I, & Castillo-Chávez, C. (2006). The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and Its Applications*, *364*, 513–536.

Bi, Z., Kan, T., Mi, C. C., Zhang, Y., Zhao, Z., & Keoleian, G. A. (2016). A review of wireless power transfer for electric vehicles: Prospects to enhance sustainable mobility. *Applied Energy*, *179*, 413–425.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*(4–5), 175–308.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, *92*(5), 1170–1182.

Bonacich, Phillip. (1991). Simultaneous group and individual centralities. *Social Networks*, *13*(2), 155–168.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, *27*(1), 55–71.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, *25*(2), 163–177.

Brandes, U., & Fleischer, D. (2005). Centrality measures based on current flow. *STACS* (vol. 3404, pp. 533–544). Berlin, Heidelberg: Springer.

Chen, J., & Safro, I. (2011). Algebraic distance on graphs. *SIAM Journal on Scientific Computing*, *33*(6), 3468–3490.

Chen, Z., He, F., & Yin, Y. (2016). Optimal deployment of charging lanes for electric vehicles in transportation networks. *Transportation Research Part B: Methodological*, *91*, 344–365.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

Cirimele, V., Freschi, F., & Guglielmi, P. (2014). Wireless power transfer structure design for electric vehicle in charge while driving. *2014 International Conference on Electrical Machines (ICEM). IEEE* (pp. 2461–2467).
Cohen, R., Erez, K., Ben-Avraham, D., & Havlin, S. (2001). Breakdown of the internet under intentional attack. *Physical Review Letters*, *86*(16), 3682.
Colizza, V., Pastor-Satorras, R., & Vespignani, A. (2007). Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics*, *3*(4), 276.
Crucitti, P., Latora, V., & Porta, S. (2006). Centrality in networks of urban streets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *16*(1), 015113.
Davis, T. A., & Hu, Y. (2011). The university of florida sparse matrix collection. *ACM Transactions on Mathematical Software (TOMS)*, *38*(1), 1.
Diekmann, O., & Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation* (Vol. 5). John Wiley & Sons.
Fouss, F., Pirotte, A., Renders, J-M., & Saerens, M. (2007). Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE Transactions on Knowledge and Data Engineering*, *19*(3), 355–369.
Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239.
Freeman, L. C, Borgatti, S. P, & White, D. R. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, *13*(2), 141–154.
Fuller, M. (2016). Wireless charging in california: Range, recharge, and vehicle electrification. *Transportation Research Part C: Emerging Technologies*, *67*, 343–356.
Ghosh, R., & Lerman, K. (2012). Rethinking centrality: The role of dynamical processes in social network analysis. *arxiv preprint arxiv:1209.4616*.
Guimerà, R., Díaz-Guilera, A., Vega-Redondo, F., Cabrales, A., & Arenas, A. (2002). Optimal network topologies for local search with congestion. *Physical Review Letters*, *89*(24), 248701.
Gutfraind, A., Safro, I., & Meyers, L. A. (2015). Multiscale network generation. *2015 18th international conference on information fusion (fusion)* (pp. 158–165). IEEE.
Hethcote, H. W. (2000). The mathematics of infectious diseases. *Siam Review*, *42*(4), 599–653.
Holme, P. (2003). Congestion and centrality in traffic flow on complex networks. *Advances in Complex Systems*, *6*(02), 163–176.
Huisinga, T., Barlovic, R., Knospe, W., Schadschneider, A., & Schreckenberg, M. (2001). A microscopic model for packet transport in the internet. *Physica A: Statistical Mechanics and Its Applications*, *294*(1–2), 249–256.
Huitema, C. (2000). *Routing in the internet*. Prentice-Hall.
Interdonato, R., & Tagarelli, A. (2016). To trust or not to trust lurkers? Evaluation of lurking and trustworthiness in ranking problems. *International conference and school on network science* (pp. 43–56). Springer.
Jang, Y. J., Ko, Y. D., & Jeong, S. (2012). Optimal design of the wireless charging electric vehicle. *2012 IEEE international electric vehicle conference (IEVC)* (pp. 1–5). IEEE.
Jayasinghe, A., Sano, K., & Nishiuchi, H. (2015). Explaining traffic flow patterns using centrality measures. *International Journal for Traffic and Transport Engineering*, *5*(2), 134–149.
Jayaweera, I. M. L. N., Perera, K. K. K. R., & Munasinghe, J. (2017). Centrality measures to identify traffic congestion on road networks: A case study of Sri Lanka. *IOSR Journal of Mathematics (IOSR-JM)*.
Jiang, B., & Claramunt, C. (2004). A structural approach to the model generalization of an urban street network. *Geoinformatica*, *8*(2), 157–171.
Jiang, B., & Jia, T. (2011). Agent-based simulation of human movement shaped by the underlying street structure. *International Journal of Geographical Information Science*, *25*(1), 51–64.
Katz, J. (1998). Luring the lurkers. *Retrieved march*, **1**(1999), 1999.
Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, *18*(1), 39–43.
Keeling, M. J, & Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton University Press.
Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*(1/2), 81–93.
Kephart, J. O, Sorkin, G. B, Chess, D. M, & White, S. R. (1997). Fighting computer viruses. *Scientific American*, *277*(5), 88–93.
Khan, M. D., Chowdhury, M., Khan, S. M., Safro, I., & Ushijima-Mwesigwa, H. (2018). *Utility maximization framework for opportunistic wireless electric vehicle charging*. Paper presented at Transportation Research Board 97th Annual Meeting, Transportation Research Board.
Khan, Z., Khan, S. M., Chowdhury, M., Safro, I., & Ushijima-Mwesigwa, H. (2019). Wireless charging utility maximization and intersection control delay minimization framework for electric vehicles. *Computer-Aided Civil and Infrastructure Engineering*.
Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., & Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature Physics*, *6*(11), 888.
Kleinfeld, J. S. (2002). The small world problem. *Society*, *39*(2), 61–66.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

Klemm, K., Serrano, M. Á., Eguíluz, V. M, & San Miguel, M. (2012). A measure of individual role in collective dynamics. *Scientific Reports*, *2*, 292.

Lai, H-M., & Chen, T. T. (2014). Knowledge sharing in interest online communities: A comparison of posters and lurkers. *Computers in Human Behavior*, *35*, 295–306.

Leskovec, J., Adamic, L. A., & Huberman, B. A. (2007a). The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, *1*(1), 5.

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007b). Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *1*(1), 2.

Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H. E., & Havlin, S. (2015). Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proceedings of the National Academy of Sciences*, *112*(3), 669–672.

Li, S., & Mi, C. C. (2015). Wireless power transfer for electric vehicle applications. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, *3*(1), 4–17.

Liu, F., Ren, Y., & Shan, X. M. (2002). A simple cellular automata model for packet transport in the internet. *Acta Physica Sinica*, *51*(6), 1175–1180.

Liu, J-G., Wu, Z-X., & Wang, F. (2007). Opinion spreading and consensus formation on square lattice. *International Journal of Modern Physics C*, *18*(07), 1087–1094.

Liu, J-G., Lin, J-H., Guo, Q., & Zhou, T. (2016). Locating influential nodes via dynamics-sensitive centrality. *Scientific Reports*, *6*, 21380.

Lukic, S., & Pantic, Z. (2013). Cutting the cord: Static and dynamic inductive wireless charging of electric vehicles. *IEEE Electrification Magazine*, *1*(1), 57–64.

Marett, K., & Joshi, K. D. (2009). The decision to share information and rumors: Examining the role of motivation in an online discussion forum. *Communications of the Association for Information Systems*, *24*(1), 4.

Mason, B. (1999). Issues in virtual ethnography. *Ethnographic studies in real and virtual environments: Inhabited information spaces and connected communities* (pp. 61–69).

Moreno, Y., Nekovee, M., & Pacheco, A. F. (2004). Dynamics of rumor spreading in complex networks. *Physical Review E*, *69*(6), 066130.

Motter, A. E. (2004). Cascade control and defense in complex networks. *Physical Review Letters*, *93*(9), 098701.

Nagel, K. (1996). Particle hopping models and traffic flow theory. *Physical Review E*, *53*(5), 4655.

Newman, M. E. J. (2003). The structure and function of complex networks. *Siam Review*, *45*(2), 167–256.

Newman, M. E. J. (2005). A measure of betweenness centrality based on random walks. *Social Networks*, *27*(1), 39–54.

Ning, P., Miller, J. M., Onar, O. C., & White, C. P. (2013). A compact wireless charging system for electric vehicles. *2013 IEEE energy conversion congress and exposition (ECCE)*. IEEE (pp. 3629–3634).

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The pagerank citation ranking: Bringing order to the web.* Tech. Rept. Stanford InfoLab.

Park, K., & Yilmaz, A. (2010). *A social network analysis approach to analyze road networks*. Paper presented at ASPRS Annual Conference. San Diego, CA.

Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, *86*(14), 3200.

Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design*, *33*(5), 705–725.

Preece, J., Nonnecke, B., & Andrews, D. (2004). The top five reasons for lurking: improving community experiences for everyone. *Computers in Human Behavior*, *20*(2), 201–223.

Qiu, C., Chau, K. T., Liu, C., & Chan, C. C. (2013). Overview of wireless power transfer for electric vehicle charging. *2013 world electric vehicle symposium and exhibition (EVS27)*. IEEE (pp. 1–9).

Riemann, R., Wang, D. Z. W., & Busch, F. (2015). Optimal location of wireless charging facilities for electric vehicles: Flow-capturing location model with stochastic user equilibrium. *Transportation Research Part C: Emerging Technologies*, *58*, 1–12.

Ripeanu, M., & Foster, I. (2002). Mapping the Gnutella network: Macroscopic properties of large-scale peer-to-peer systems. *International workshop on peer-to-peer systems* (pp. 85–93). Springer.

Rossi, R., & Ahmed, N. (2015). The network data repository with interactive graph analytics and visualization. *AAAI* (vol. 15, pp. 4292–4293).

Scheurer, J., Curtis, C., & Porta, S. (2008). *Spatial network analysis of multimodal transport systems: Developing a strategic planning tool to assess the congruence of movement and urban structure: a case study of perth before and after the perth-to-mandurah railway*. Paper presented at GAMUT, Australasian Centre for the Governance and Management of Urban Transport, University of Melbourne.

Schlosser, A. E. (2005). Posting versus lurking: Communicating in a multiple audience context. *Journal of Consumer Research*, *32*(2), 260–265.

Shaydulin, R., Chen, J., & Safro, I. (2017). Relaxation-based coarsening for multilevel hypergraph partitioning. *SIAM Multiscale Modeling and Simulation*, preprint arxiv:1710.06552.

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press

Šikić, M., Lančić, A., Antulov-Fantulin, N., & Štefančić, H. (2013). Epidemic centralities there an underestimated epidemic impact of network peripheral nodes? *The European Physical Journal B*, *86*(10), 440.

Soroka, V., Jacovi, M., & Ur, S. (2003). We can see you: a study of communities invisible people through reachout. *Communities and technologies* (pp. 65–79). Springer.

Southworth, M., & Ben-Joseph, E. (2013). *Streets and the shaping of towns and cities*. Island Press.

Spring, N., Mahajan, R., & Wetherall, D. (2002). Measuring ISP topologies with rocketfuel. *SIGCOMM* (vol. 32, pp. 133–145).

Staudt, C. L., Hamann, M., Gutfraind, A., Safro, I., & Meyerhenke, H. (2017). Generating realistic scaled complex networks. *Applied Network Science*, *2*(1), 36.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268.

Tagarelli, A., & Interdonato, R. (2013). Who's out there? Identifying and ranking lurkers in social networks. *Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining* (pp. 215–222). ACM.

Tagarelli, A., & Interdonato, R. (2014). Lurking in social networks: topology-based analysis and ranking methods. *Social Network Analysis and Mining*, *4*(1), 230.

Tao, Z., Zhongqian, F., & Binghong, W. (2006). Epidemic dynamics on complex networks. *Progress in Natural Science*, *16*(5), 452–457.

Tsoumakos, D., & Roussopoulos, N. (2006). Analysis and comparison of p2p search methods. *Proceedings of the 1st international conference on scalable information systems* (p. 25). ACM.

Ushijima-Mwesigwa, H., Khan, M. D., Chowdhury, M. A., & Safro, I. (2017). Optimal installation for electric vehicle wireless charging lanes. *arxiv preprint arxiv:1704.01022*.

Van Mierlo, T. (2014). The 1% rule in four digital health social networks: An observational study. *Journal of Medical Internet Research*, *16*(2).

Vilathgamuwa, D. M., & Sampath, J. P. K. (2015). Wireless power transfer for electric vehicles, present and future trends. *Plug in electric vehicles in smart grids* (pp. 33–60). Springer.

Wang, P., Hunter, T., Bayen, A. M., Schechtner, K., & González, M. C. (2012). Understanding road usage patterns in urban areas. *Scientific Reports*, *2*, 1001.

Wang, Y., Yun, X., & Li, Y. (2007). Analyzing the characteristics of gnutella overlays. *Fourth international conference on information technology, 2007 (ITNG'07)* (pp. 1095–1100). IEEE.

Watts, D. J., Peretti, J., & Frumin, M. (2007). *Viral marketing for the real world*. Harvard Business School Pub.

Wood, R. K. (1993). Deterministic network interdiction. *Mathematical and Computer Modelling*, *17*(2), 1–18.

Yan, G., Zhou, T., Hu, B., Fu, Z-Q., & Wang, B-H. (2006). Efficient routing on complex networks. *Physical Review E*, *73*(4), 046108.

Zhang, Y., Wang, X., Zeng, P., & Chen, X. (2011). Centrality characteristics of road network patterns of traffic analysis zones. *Transportation Research Record: Journal of the Transportation Research Board*, *2256*(1), 16–24.

Zhao, L., Xie, W., Gao, H. O., Qiu, X., Wang, X., & Zhang, S. (2013). A rumor spreading model with variable forgetting rate. *Physica A: Statistical Mechanics and Its Applications*, *392*(23), 6146–6154.

Cite this article: Ushijima-Mwesigwa H., Khan Z., Chowdhury M. A., and Safro I. (2019). Centralities for networks with consumable resources. *Network Science* **7**, 376–401. https://doi.org/10.1017/nws.2019.7

https://doi.org/10.1017/nws.2019.7 Published online by Cambridge University Press