# Review of "Communicability Across Evolving Networks by Grindrod, P., Higham, D. J., Parsons M. C. and Estrada, E."

Kanak Choudhury
Department of Statistics
University of Central Florida
Orlando, FL, USA
Email: kanakc@knights.ucf.edu

## I. Introduction

CONNECTEDNESSConnectedness, path length, diameter, degree and clique play the main important factors to establish the network science. In this article, authors tried to establish a new type of time-dependent network model. In Figure 1, it was represented the asymmetry network caused by the arrow of time although each individual network is symmetric. That was the motivation to distinguish between static and dynamic networks and pointed out the need for a theory that can deal with-

1) the time ordering inherent in the edge lists when considering communication around the network,
2) Respects the inherent asymmetry imposed by the arrow of time, even when each individual snapshot consists of an undirected network [1].

There are many applications where the network connectivity patterns changes over time. Examples include network of instant messaging systems (Facebook, MSN) [2], [3]; networks of travelers, vehicles dened over a dynamic transportation infrastructure [4]–[6]; correlated neural activity in response to a functional task [7].

In this article, authors implemented a new extended centrality concept, that was related to [8]–[10], to address dynamic network. Authors tried to emphasize in this article that in the time-dependent network, the population of nodes remained constant from the outset and the graph evolves through the appearance (birth) or the deletion (death) of edges.

## II. Katz Centrality

In this section, authors described Katz centrality [11] that works with a single, static network. Let $A$ denote $N$-by-$N$ binary adjacency matrix for a directed graph $G$ defined over $N$ nodes, where $A_{ij}$ represents one if there is a link from node $i$ to node $j$ and $A_{ij} \neq A_{ji}$ so that the adjacency matrix may be asymmetric.

To find the propensity for node $i$ to communicate, or interact, with node $j$, it was counted the number of walks of length $w$ from $i$ to $j$ and combined these counts into a single, cumulative total over all $w$. The walks of length $w$ were scaled by a factor $a^w$ to find the shorter walks, where $a$ is a suitably chosen scalar. The $k$th power of the adjacency matrix has $i$, $j$ element that counts the number of walks of length $w$ from node $i$ to node $j$ and the expansion of $I + aA + a^2A^2 + a^3A^3 + \cdots$ converges to $(I - aA)^{-1}$ where $I \in \mathbb{R}^{N\times N}$, $a < 1/\rho(A)$ , and $\rho(\cdot)$ denotes the largest absolute eigenvalue. Since $((I - aA)^{-1})_{ij}$ summarizes how well information can pass from node $i$ to node $j$ and the $n$th row sum

$$\sum_{k=1}^{N} \left((I - aA)^{-1}\right)_{nk} \tag{1}$$

represents the Katz centrality. This centrality measure is based on the combinatorics of walks, which allow nodes and edges to be reused during a traversal, rather than paths or shortest paths. The Katz centrality (1) measures the ability of node $n$ to send out information along the directed links. On the other hand, to measure the ability of node $n$ to acquire information can be found by column sum of $((I - aA)^{-1})_{ij}$:

$$\sum_{k=1}^{N} \left((I - aA)^{-1}\right)_{kn}.$$

## III. Dynamic Centralities

In this section, authors described their proposed method of dynamic centrality for network. It was considered, for a given set of $N$ nodes, an ordered sequence $G^{[k]}$ for $k = 0, 1, 2, \cdots, M$ with time points $t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_M$, where each $G^{[k]}$ is an unweighted graph defined over those nodes and time $t_k$ with corresponding adjacency matrix, $A^{[k]}$. The generalization of adjacency matrix, $A^{[k]}$ for the static graph concept of a walk was defined to explain how well information can be passed between pairs of nodes as following.

*Definition 1*: A dynamic walk of length $w$ from node $i_1$ to node $i_{w+1}$ consists of a sequence of edges $i_1 \rightarrow i_2, i_2 \rightarrow i_3, \cdots, i_w \rightarrow i_{w+1}$ and a non-decreasing sequence of times $t_{r_1} \leq t_{r_2} \leq t_{r_3} \leq \cdots \leq t_{r_w}$ such that $A^{[r_m]}_{i_m,i_{m+1}} \neq 0$ and the lifetime of this walk to be $t_{r_w} - t_{r_1}$ [1].

Note that, the sequence of times $t_{r_1}, t_{r_2}, \cdots, t_{r_w}$ must be non-decreasing but it is possible to consider repeated times, i.e. $r_1 < r_2 = r_3 < r_4$ meaning two edges are followed at

Department of Statistics, University of Central Florida, FL, USA

time $t_{r_2}$, and not required consecutive times, i.e. if $r_2 > r_1+1$ then there is no walk between $t_{r_1}$ and $t_{r_2}$ times.

In this dynamic setting, authors used the same concept that were used to derive the Katz centrality measure (1). Using down-weighting walks of length $w$ by a factor $a^w$ on the dynamic sequence, the following matrix product was obtained.

$$Q := (I - aA^{[0]})^{-1}(I - aA^{[1]})^{-1} \cdots (I - aA^{[M]})^{-1} \quad (2)$$

where, $a < {}^{1}/_{\max_s \rho(A^{[s]})}$. Note that, the identity matrices in (2) explained the concept of waiting time of a message at a node until a suitable connection appears at a later time.

So, the centrality measures for the $n$th node were obtained from $Q_{ij}$, which explains how well information can be passed from node $i$ to node $j$, by row and column sums

$$C_n^(broadcast) := \sum_{k=1}^{N} Q_{nk} \text{ and } C_n^(receive) := \sum_{k=1}^{N} Q_{kn} \quad (3)$$

and represented by how effectively node $n$ can broadcast and receive messages, respectively.

The normalized $Q$, to avoid under or overflow and to have relative values of the centrality measures across all nodes, can be obtained by

$$\hat{Q}^{[k]} = \frac{\hat{Q}^{[k-1]}(I - aA^{[k]})^{-1})}{\left\|\hat{Q}^{[k-1]}(I - aA^{[k]})^{-1}\right\|}, k = 0, 1, 2, , M$$

where, $\hat{Q}^{[-1]} = I$. This technique bears two features- (i) computations are solution of linear system equations which is convenient and efficient for large, sparse networks, and (ii) the inherent asymmetry caused by the dynamics is captured directly through the non-commutativity of matrix multiplication.

## IV. Computational Tests

In this section, authors illustrated the new dynamic centralities with some simulated and real data.

### A. Synthetic Data

To compare the dynamic centrality (3) with Katz centrality (1), simulated networks were generated for $N = 1001$ nodes with 31 time points. In this simulated networks 1000 nodes were constructed independently and node 1001 was connected to the two nodes with largest degree. Figure 2 showed that dynamic centrality correctly identified the importance of node 1001 that was absent for the Katz centrality for different parameters.

### B. Telecommunication Data

Here telecommunication data [12] was used to illustrate the new dynamic centrality technique. Figure 3 showed a summary of the adjacency matrices aggregated into 28 day intervals. Figure 4 showed the daily edge count, and centrality measures with respect to broadcast, receiver and total degree. Figure 5 examined how the centralities changes to the change of parameter $a$.

### C. Email Data

The new dynamic centrality technique was also illustrated on a public domain data set concerning email activities of Enron employees [13]. Figure 6 showed the daily edge count, scatter plots with respect to broadcast vs receiver centralities, broadcast vs total out degree and receive vs total in degree. It was found that the two new centrality measures were distinct; in particular, only two nodes appeared in the overlap of top twenty broadcast and receive and it was clear that some top receivers were very poor broadcasters. Figure 7 showed how the new centralities change with changes of $a$, indicated robustness in this parameter regime and also showed the eect of symmetrizing the data.

## V. Discussion

The new centrality measures proposed in this article can be used to monitor network behavior dynamically. Based on the previous data, it is possible to obtain the expected future communicability using Markovian concept. The expected value of the future adjacency matrix for a given $H_p$, where $H_p = A^{[p]}, A^{[p-1]}, \cdots$, can be obtained by $E(A^{[p']} \mid H_p)$, where $p' > p$. Then the estimated expectation of the communicability over the current and future time steps can be obtained by-

$$E(S(Q) \mid H_0) = I + a \sum_{p=0}^{M} E(A^{[p']} \mid H_0) + O(a^2)$$

*and*

$$E(AS(Q) \mid H_0) = a^2 \sum_{p=0}^{M} \sum_{p'=p+1}^{M} E([A^{[p]}, A^{[p']}] \mid H_0) + O(a^3)$$

For $p \to \infty$, it is found that $A_{ij}^{[\infty]} = {}^{\alpha_{ij}}/_{(\alpha_{ij}+w_{ij})}$, where $\alpha_{ij}$ and $w_{ij}$ denoted as the stepwise *birth* and *death* rates. Then, for considering time steps 0 to $M$, it was found that

$$E(Q \mid A^{[0]}) = I + a(R_M \circ (A^{[0]} - A^{[\infty]}) + (M+1)A^{[\infty]} + O(a^2)$$

where $R_M$ is the symmetric matrix given by $(R_p)_{ij} = {}^{(1-(1-\alpha_{ij}+w_{ij})^{M+1})}/_{(\alpha_{ij}+w_{ij})}$ and $\circ$ denotes component-wise multiplication. This is nothing but the limiting value of the Markov chain that explains the relative contributions to $Q$ made by the initial condition and the long term expected equilibrium value for each edge.

So if the observer wants to take some action on the future state, it is possible to predict long term expected communication based on the current state of network.

## References


[1] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, "Communicability across evolving networks," *Physical Review E*, vol. 83, no. 4, p. 046120, 2011.

[2] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora, "Small-world behavior in time-varying graphs," *Physical Review E*, vol. 81, no. 5, p. 055101, 2010.

[3] J. Leskovec and E. Horvitz, "Planetary-scale views on a large instant-messaging network," in *Proceedings of the 17th international conference on World Wide Web*, pp. 915–924, ACM, 2008.

[4] A. Gautreau, A. Barrat, and M. Barthélemy, "Microdynamics in stationary complex networks," *Proceedings of the National Academy of Sciences*, vol. 106, no. 22, pp. 8847–8852, 2009.

[5] K. A. Berman, "Vulnerability of scheduled networks and a generalization of menger's theorem," *Networks*, vol. 28, no. 3, pp. 125–134, 1996.

[6] L. McNamara, C. Mascolo, and L. Capra, "Media sharing based on colocation prediction in urban transport," in *Proceedings of the 14th ACM international conference on Mobile computing and networking*, pp. 58–69, ACM, 2008.

[7] P. Grindrod and D. J. Higham, "Evolving graphs: dynamical models, inverse problems and propagation," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 466, pp. 753–770, The Royal Society, 2010.

[8] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, "Characterising temporal distance and reachability in mobile and online social networks," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, pp. 118–124, 2010.

[9] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, "Temporal distance metrics for social network analysis," in *Proceedings of the 2nd ACM workshop on Online social networks*, pp. 31–36, ACM, 2009.

[10] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia, "Analysing information flows and key mediators through temporal centrality metrics," in *Proceedings of the 3rd Workshop on Social Network Systems*, p. 3, ACM, 2010.

[11] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.

[12] N. Eagle, A. S. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the national academy of sciences*, vol. 106, no. 36, pp. 15274–15278, 2009.

[13] A. Chapanond, M. S. Krishnamoorthy, and B. Yener, "Graph theoretic and spectral analysis of enron email data," *Computational & Mathematical Organization Theory*, vol. 11, no. 3, pp. 265–281, 2005.