Nonbacktracking Walk Centrality for Directed Networks

F. Arrigo, P. Grindrod, D. J. Higham, and V. Noferini

27th Biennial NA Conference
University of Strathclyde
June 27th–30th, 2017
Complex Networks

Walk-based centrality measures

Nonbacktracking walks

Numerical example

Summing up

Figure from: http://www.npr.org/2016/04/16/474396452/how-math-determines-the-game-of-thrones-protagonist
Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an unweighted complex network with $n$ nodes. Its adjacency matrix is $A = (a_{ij}) \in \mathbb{R}^{n \times n}$:

$$a_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

An edge $(i, j) \in E$ s.t. $(j, i) \in E$ is called reciprocal.
A walk of length $r$ is a sequence of $r + 1$ nodes

$$i_1, i_2, \ldots, i_{r+1}$$

such that $(i_l, i_{l+1}) \in E$ for all $1 \leq l \leq r$.

The walk is said to be closed if $i_1 = i_{r+1}$.

The quantities $(A^r)_{ii}$, $(A^r)_{ij}$ count closed (resp., open) walks of length $r$.

Example: the degree of node $i$ is defined as

$$d_i = |\{ j \in V : (i, j) \in E \}|$$

$$= (A^2)_{ii} = \sum_{j=1}^{n} a_{ij}a_{ji}.$$
Walk-based centrality measures
A node centrality measure is a function

\[ m : \mathcal{V} \rightarrow \mathbb{R}_{\geq 0} \]

which is invariant under graph isomorphism\(^1\) and that assigns to each node in the graph a nonnegative score that quantifies its importance within the network.

---

\(^1\) Two graphs \(G_1\) and \(G_2\) with associated adjacency matrices \(A_1\) and \(A_2\) are isomorphic if there exists a permutation matrix \(P\) such that \(A_1 = PA_2P^T\).
The degree centrality

\[ d_i = e_i^T A_1 = \sum_{j=1}^{n} a_{ij}. \]

It leads to \( d = A_1 \).

The degree centrality is “too local”.
The **degree centrality**

\[ d_i = \mathbf{e}_i^T \mathbf{A} \mathbf{1} = \sum_{j=1}^{n} a_{ij}. \]

It leads to \( d = \mathbf{A} \mathbf{1} \).

The degree centrality is “too local”.

Bonacich introduced the **eigenvector centrality**:

\[ x_i \propto \sum_{j=1}^{n} a_{ij} x_j. \]

It leads to \( \mathbf{A} \mathbf{x} = \lambda \mathbf{x} \) and, if \( \mathbf{A} \) is irreducible, then \( \mathbf{x} \) is the **Perron vector** of \( \mathbf{A} \) and \( \lambda = \rho(\mathbf{A}) \).
Katz Centrality

Let \( f(x) = \sum_{r=0}^{\infty} c_r x^r \), then, within the radius of convergence:

\[
f(A) = \sum_{r=0}^{\infty} c_r A^r
\]

Looking closely at \( (f(A))_{ij} \):

- it tells us how many walks (up to infinite length) originate at node \( i \) and end at node \( j \)
- if \( c_r \geq 0 \) and \( c_r \to 0 \) as \( r \to \infty \), longer walks are given less importance.

Katz centrality:

\[
k = 1 + \left( \sum_{r=1}^{\infty} \alpha^r A^r \right) 1 = (I - \alpha A)^{-1} 1,
\]

where \( \alpha \in (0, 1/\rho(A)) \).

\( \Rightarrow \) solve a \text{ sparse} linear system.
Nonbacktracking walks
A walk is said to be \textit{backtracking} (BTW) if it contains at least one sequence of nodes of the form 
\[ i \ell i, \]
\textit{nonbacktracking} (NBTW) otherwise.
Let $p_r(A) \in \mathbb{R}^{n \times n}$ be such that

$$(p_r(A))_{ij} = |\{\text{NBTW s of length } r \text{ from node } i \text{ to node } j\}|.$$

It is the nonbacktracking analogue of the matrix power $A^r$.

We define a **NBTW-based centrality** measure as:

$$b = 1 + \left( \sum_{r=1}^{\infty} t^r p_r(A) \right) \mathbf{1} = \phi(A, t) \mathbf{1},$$

where $t > 0$ is chosen so that $\phi(A, t) = \sum_r t^r p_r(A)$ converges.

**Questions:**

- is this feasible?
- restrictions on $t$?
Theorem: Let $A$ be the adjacency matrix of a digraph, $D = \text{diag}(\text{diag}(A^2))$, and $S = A \circ A^T$. Then,

\begin{align*}
p_0(A) &= I, \\
p_1(A) &= A, \\
p_2(A) &= A^2 - D
\end{align*}

and for all $r \geq 3$

\begin{align*}
p_r(A) &= A p_{r-1}(A) + (I - D) p_{r-2}(A) - (A - S) p_{r-3}(A).
\end{align*}

Theorem Let $A$, $S$, and $D$ be defined as before. Moreover, let

$$\phi(A, t) = \sum_{r=0}^{\infty} t^r p_r(A)$$

and

$$M(t) = I - At + (D - I)t^2 + (A - S)t^3.$$ 

Then

$$M(t) \phi(A, t) = (1 - t^2)I$$
Recalling that $b = \phi(A, t) \mathbf{1}$, we can thus write

$$(I - tA)b = (1 - t^2)\mathbf{1},$$

where

$$A := [A + (I - D)t + (S - A)t^2].$$
Recalling that \( b = \phi(A, t)1 \), we can thus write
\[
(l - tA)b = (1 - t^2)1,
\]
where
\[
A := [A + (l - D)t + (S - A)t^2].
\]

The NBTW-based centrality measure is then:
\[
b = (1 - t^2)(l - tA)^{-1}1,
\]
where \( t > 0 \) is chosen so that \( \phi(A, t) = \sum_r t^r p_r(A) \) converges.

\[\Rightarrow\text{ same cost as Katz.}\]
Original graph: A

\[
\begin{array}{c}
i \\
\ell \\
j
\end{array}
\]
Original graph: $A$

New graph: $A = [A + (I - D)t + (S - A)t^2]$

New graph: 

$A = [A + (I - D)t + (S - A)t^2]$
The power series $\phi(A, t) = \sum_{r=0}^{\infty} t^r p_r(A)$ converges if $0 < t < 1/\rho(C)$ where $\rho(C)$ is the spectral radius of the matrix

$$C = \begin{bmatrix} A & (I - D) & (S - A) \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix}.$$
The radius of convergence $R_\phi$ of the power series $\phi(A, t) = \sum_{r=0}^{\infty} t^r p_r(A)$ is

$$R_\phi = \begin{cases} 
1/\rho(C) & \text{if there is at least one directed cycle} \\
\infty & \text{if there are no directed cycles}
\end{cases}$$

In the latter case, $\rho(C) = 1.$
The radius of convergence $R_{\phi}$ of the power series $\phi(A, t) = \sum_{r=0}^{\infty} t^r p_r(A)$ is

$$R_{\phi} = \begin{cases} 
\frac{1}{\rho(C)} & \text{if there is at least one directed cycle} \\
\infty & \text{if there are no directed cycles}
\end{cases}$$

In the latter case, $\rho(C) = 1$.

Let $R_\psi$ be the radius of convergence of $\psi(A, t) = \sum_{r=0}^{\infty} t^r A^r$ and let $R_{\phi}$ be the radius of convergence of $\phi(A, t) = \sum_{r=0}^{\infty} t^r p_r(A)$. Then, $R_\psi \leq R_{\phi}$. 
Let $t \in (0, \rho(C)^{-1})$. Then the NBTW centrality vector $b(t)$ returns the same ranking as that returned by

- $d^\text{out} = A1$ as $t \to 0^+$
- the first $n$ components of $x$: $Cx = \rho(C)x$, if the rank of $(I - \rho(C)^{-1}C)$ is $3n - 1$ and $t \to (1/\rho(C))^{-}$.

This theorem generalizes:

Numerical example
The network PAJEK/GLOSSGT represents connections between words from the graph/digraph glossary.

- largest weakly connected component contains 60 nodes;
- $\rho(A) = \rho(C) = 1$;
- we use $\alpha = t = 0.9$. 
Katz vs. NBTW

Complex Networks
Walk-based centrality measures
Nonbacktracking walks
Numerical example
Summing up
Katz vs. NBTW

Complex Networks
Walk-based centrality measures
Nonbacktracking walks
Numerical example
Summing up
Katz vs. NBTW

Complex Networks
Walk-based centrality measures
Nonbacktracking walks
Numerical example
Summing up
Summing up
We worked on **directed, static** networks and considered the non-backtracking version of **“standard” walks**.

We did:

- define a NBTW-based centrality measure
- demonstrated computational feasibility
- characterized convergence
- generalized eigenvector centrality of Martin et al.
- **In the paper:** showed that pruning certain nodes adds to efficiency
- **In the paper:** showed that NBTW Katz dampen localization effects

**What are we doing now?** We are considering other types of...