Title

Two steps to obfuscation

Journal

Social Networks, 39(1)

ISSN

0378-8733

Authors

Friedkin, Noah E
Johnsen, Eugene C

Publication Date

2014-10-01

DOI

10.1016/j.socnet.2014.03.008

Peer reviewed

Social Networks 39 (2014) 12–13

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Two steps to obfuscation

Noah E. Friedkin[a,b,*], Eugene C. Johnsen[c]

[a] Department of Sociology, University of California, Santa Barbara, United States
[b] Center for Control, Dynamical Systems and Computation, College of Engineering, University of California, Santa Barbara, United States
[c] Department of Mathematics, University of California, Santa Barbara, United States

ABSTRACT

This note addresses the historical antecedents of the 1998 PageRank measure of centrality. An identity relation links it to the 1990–1991 models of Friedkin and Johnsen.

1. Introduction

Friedkin and Johnsen (1990) presented a model of a multi-agent network in which the total influences of the agents are related to the number and length of the walks in the network as follows:

$$\mathbf{V} = (\mathbf{I} + \alpha\mathbf{W} + \alpha^2\mathbf{W}^2 + \alpha^3\mathbf{W}^3 + \cdots)(1-\alpha) = (\mathbf{I} - \alpha\mathbf{W})^{-1}(1-\alpha), \quad (1)$$

where $\mathbf{W}_{n\times n}$ is the row-stochastic matrix associated with the influence network of agents' relative direct allocated weights and $0<\alpha<1$ is a scalar corresponding to the relative contributions of the influence network and the initial state-space conditions of the agents. In applications to a discrete-time opinion formation process, the factor $1-\alpha$ corresponds to the weight agents assign to their initial positions (and all conditions that generated their initial positions) during each period of the process. While $(\mathbf{I}-\alpha\mathbf{W})^{-1}$ is a familiar construct in network analysis (Katz, 1953; Hubbell, 1965), the $\mathbf{V}$ construct and its row-stochasticity appear as derived implications of a postulated second-order convex combination influence mechanism. This process-oriented formulation was novel, and seminal to subsequent work in which the homogeneity of α was relaxed (Friedkin, 1998; Friedkin and Johnsen, 1999, 2011). The v_{ij} of the $n\times n$ matrix $\mathbf{V}=[v_{ij}]$ corresponds to the relative net influence of agent *j* on agent *i*.

Friedkin (1991) developed the employment of $\mathbf{V}$ as a measure of structural centrality

$$\mathbf{c} = \frac{1}{n}\mathbf{V}^T\mathbf{e}, \quad \mathbf{c}^T\mathbf{e} = 1, \quad (2)$$

$$\mathbf{c} = \left(\frac{1-\alpha}{n}\right)(\mathbf{I}-\alpha\mathbf{W}^{\mathrm{T}})^{-1}\mathbf{e}, \quad (3)$$

where here, and henceforth, $\mathbf{e}$ is a $n\times 1$ vector of ones and each element of the $n\times 1$ vector $\mathbf{c}=[c_i]$ is "the average total effect centrality of an actor" (Friedkin, 1991, pp. 1485–1487). The average may be based on the n values of each column, or $n-1$ values when the main diagonal values of $\mathbf{V}$ are excluded. The latter concentrates the measure on the total effects of an agent *i* on *other* agents.

If the vector of averages in Eq. (2) are expressed as follows

$$\mathbf{c} = \left(\frac{1-\alpha}{n}\right)\mathbf{e} + \alpha\mathbf{W}^T\mathbf{c}, \quad (4)$$

then Eq. (3) are their solutions. If no averages are taken, then the model simply presents the sums of the columns of $\mathbf{V}$

$$\mathbf{c} = \mathbf{V}^T\mathbf{e} = (1-\alpha)(\mathbf{I}-\alpha\mathrm{W}^{\mathrm{T}})^{-1}\mathbf{e}, \quad (5)$$

whence

$$\mathbf{c} = (1-\alpha)\mathbf{e} + \alpha\mathbf{W}^T\mathbf{c}. \quad (6)$$

We will now show why this odd form of the model (Eq. (6)) is of interest.

2. Two steps to obfuscation

Consider an application of the model to a webgraph composed of nodes that are the pages of the webgraph and edges that are its hyperlinks. Let $\mathbf{A}=[a_{ij}]$ be the $n\times n$ adjacency matrix of the webgraph, where $a_{ij}=1$ if page *i* has a directed link to page *j* and 0 otherwise. Let $\mathbf{W}=[w_{ij}]$ be the $n\times n$ normalized adjacency matrix,

$$w_{ij} = \frac{a_{ij}}{\sum_{k=1}^n a_{ik}} = \frac{a_{ij}}{od(i)}, \quad od(i) > 0, \quad (7)$$

* Corresponding author at: Department of Sociology, University of California, Santa Barbara, United States. Tel.: +1 805 259 8389.
E-mail address: friedkin@soc.ucsb.edu (N.E. Friedkin).

http://dx.doi.org/10.1016/j.socnet.2014.03.008
0378-8733/

for all i and j. Eq. (6) may now be expressed as follows

$$c_i = (1-\alpha) + \alpha \sum_{j \in S}^{n} \frac{c_j}{od(j)}, \quad od(j) > 0, \tag{8}$$

for all i, where S is the set of edges for which j has a direct link to i.

Step 1. Now alter the notation. Let $PR(j) \equiv c_j$, $j = 1, \ldots, n$ and let $d \equiv \alpha$. Those changes of notation present

$$PR(i) = (1-d) + d \sum_{j \in S}^{n} \frac{PR(j)}{od(j)}, \quad od(j) > 0. \tag{9}$$

Step 2. Now alter the remaining notation. Let A be i and let $T1, \ldots, Tn$ be the $j = 1, \ldots, n$ pages that point to it. Let $C(A)$ be the number of links going out of page A. Those changes of notation present a scalar equation for each page A of the webgraph

$$PR(A) = (1-d) + d\left[\frac{PR(T1)}{C(T1)} + \cdots + \frac{PR(Tn)}{C(Tn)}\right]. \tag{10}$$

which is exactly the description of the PageRank calculation that Brin and Page (1998) presented as the foundation of Google. The set of these scalar equations is equivalent to Eq. (6), and Eq. (5) is its solution, i.e., the unnormalized measure of centrality. However, the Page and Brin presentation of it generated some confusion and it was subsequently modified to the normalized measure

$$PR(A) = \left(\frac{1-d}{n}\right) + d\left[\frac{PR(T1)}{C(T1)} + \cdots + \frac{PR(Tn)}{C(Tn)}\right]. \tag{11}$$

This equation is equivalent to Eq. (4), and Eq. (3) is its solution. In either case, Eq. (1) provides the foundation of an algorithmic approximation of **V** when inverse computations are not feasible.

These two sets of notation have completely obscured the equivalence of the PageRank calculations and the constructs of Friedkin and Johnsen (1990) and Friedkin (1991). Mathematica now presents a PageRank centrality solution for the adjacency matrices of digraphs that, with $\alpha = 0.85$, returns centrality scores that are identical to those of Eq. (3).

3. Discussion

The PageRank formula was published over 15 years ago in a venue outside of sociology. This is "ancient" history. But it is a history that remains relevant today whenever the solution of the formula is employed as a measure of centrality. Reinvention of the wheel in the field of social networks is not an unfamiliar event now that investigators from the natural and engineering sciences have become more interested in social networks. Notation differences, and the difficulty of monitoring publications appearing in journals outside one's own discipline, have obscured the correspondence of the 1998 PageRank measure and the 1991 measure proposed by Friedkin.

Brin and Page presented the formula as an intuitive hop to a novel eigenvector-like measure. In contrast, Friedkin's (1991) measure was a development of Friedkin and Johnsen's (1990) model. The 1990 construct **V** was an analytically derived corollary of their specification of a proposed convex combination mechanism of influence among agents joined in a multi-agent network. The 1990 model, in turn, was developed as a generalization of the seminal work of French (1956). An eigenvector approach to centrality is natural and appealing. The relaxed eigenvector-like formula of Eqs. (4) and (6), and their PageRank equivalents, do not appear in our work as an intuitive hop.

References

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Comput. Netw. ISDN Syst. 30, 107–117.

French Jr., J.R.P., 1956. A formal theory of social power. Psychol. Rev. 63, 181–194.

Friedkin, N.E., 1991. Theoretical foundations for centrality measures. Am. J. Sociol. 96, 1478–1504.

Friedkin, N.E., 1998. A Structural Theory of Social Influence. Cambridge University Press, New York.

Friedkin, N.E., Johnsen, E.C., 1990. Social influence and opinions. J. Math. Sociol. 15, 193–206.

Friedkin, N.E., Johnsen, E.C., 1999. Social influence networks and opinion change. Adv. Group Process. 16, 1–29.

Friedkin, N.E., Johnsen, E.C., 2011. Social Influence Network Theory. Cambridge University Press, New York.

Hubbell, C.H., 1965. An input–output approach to clique identification. Sociometry 28, 377–399.

Katz, L., 1953. A new status index derived from sociometric analysis. Psychometrika 18, 39–43.