# Finding Spread Blockers in Dynamic Networks

Habiba[*]
Dept. Computer Science
University of Illinois at Chicago
Chicago, IL 60607, USA
hhabiba@cs.uic.edu

Yintao Yu[†]
Dept. Electronic Engineering
Shanghai Jiao Tong University
Shanghai 200240, China
yintaoyu@gmail.com

Tanya Y. Berger-Wolf[‡]
Dept. Computer Science
University of Illinois at Chicago
Chicago, IL 60607, USA
tanyabw@uic.edu

Jared Saia[§]
Dept. of Computer Science
University of New Mexico
Albuquerque, NM 87131, USA
saia@cs.unm.edu

## ABSTRACT

We study the problem of identifying nodes in a network that are good *blockers*: individuals that are most effective in blocking the spread of a dynamic process through the population. We study several structural measures for identifying the best blockers in both static and dynamic social networks. We find that for both static and dynamic networks, simple local measures such as a node's degree are surprisingly good indicators of how good a blocker it is.

## 1. INTRODUCTION

How can we stop the spread of a dynamic process through a social network? This problem has applications to many diverse areas such as preventing or inhibiting the spread of diseases [5, 15, 24], computer viruses, rumors, and undesirable fads or behaviors [12, 13, 21, 22]. A common approach to spread inhibition is to identify key individuals whose removal will most dampen the spread. In the context of the spread of a disease, it is a question of finding individuals to be quarantined or vaccinated so that the disease is prevented from becoming an epidemic. We call this set of key individuals the *blockers* of the spreading process.

There has been significant previous work related to studying and controlling the spread of dynamic processes in a network (see e.g. [27, 15, 11, 35, 28]). Unfortunately, many past results focus on either: 1) effectively starting the spread of a process rather than blocking it; or 2) identifying locations to place sensors in order to quickly determine where and when a spread has started. Moreover, algorithms proposed in previous work require computationally intensive calculations to identify critical individuals. Finally, previous work elides the dynamic nature of social interactions. The flow of information through a social network is dependent on who has the information at what point in time and who are the individuals in contact at that moment with the information carrier that are likely to acquire the information next.

In this paper, we show that simple local measures, such as degree, can effectively predict an individual's capacity to block spreading processes, even in dynamic social networks. The implication of our results are that there are practical scalable heuristics for identifying quarantine and vaccination targets in order to prevent an epidemic.

**Related Work**: Several previous results have addressed the problem of identifying the best blockers in a network [3, 15, 24, 27, 28]. Aspnes at al. [3] proposed a game theoretic approach for inoculation strategies based on sum-of-squares partition problem. In [28], Leskovec et al. propose an elegant but computationally intensive algorithm that unfortunately, works only for a stable, relatively unchanging network. Eubank et al. [15] and Kempe et al. [27] show that a variant of this problem is NP-Hard, with no proposed algorithmic solution.

## 2. DEFINITIONS

Populations of individuals interacting over time are often represented as networks, where the nodes correspond to individuals and a pairwise interaction is represented as an edge between the corresponding individuals. Typically, there is a single network representing all interactions that have happened during the entire observation period. We call this representation an *aggregate network*. We will represent aggregate networks as multigraphs: multiple interactions between a pair of individuals are represented as multiple unweighted edges between them. In this paper we also use an explicitly *dynamic network* representation (described below) that takes the history of interactions into account.

**Dynamic Network:** We represent a dynamic network as a series $\langle G_1, \ldots, G_T \rangle$ of static networks where each $G_t$ is a snapshot of individuals and their interactions at time $t$. For simplicity, we also assume that the time period is divided

---

[*] Work supported in part by the Fulbright fellowship.

[†] Work performed in part while being a visiting student at the Univeristy of New Mexico.

[‡] Work supported in part by the NSF grant IIS-0705822 and NSF CAREER Award 0747369.

[§] Work supported in part by the NSF grant IIS-0705822, NSF CAREER Award 0644058, and an AFO MURI award.

*The 2nd SNA-KDD Workshop '08 (SNA-KDD'08),* August 24, 2008, Las Vegas, Nevada, USA

into discrete steps $\{1, \dots, T\}$.

Definition 1. *[23] Let $\{1, \dots, T\}$ be a finite set of discrete timesteps. Let $V = \{1, \dots, n\}$ be a set of individuals. Let $G_t = (V_t, E_t)$ be a graph representing a snapshot of a static network at time $t$. $V_t \subseteq V$, is a subset of individuals $V$ observed at time $t$. An edge $(u_t, v_t) \in E_t$ if individuals $u$ and $v$ have interacted at time $t$. A dynamic network $G = \langle G_1, \dots, G_T \rangle$ is the graph $G = (V, E)$ of the time series of graphs $G_t$ such that $V = \bigcup_t V_t$ and $E = \bigcup_t E_t \cup \bigcup_{t-1} (v_t, v_{t+1})$.*

In general, in a dynamic network both edges and individuals can change over time. The edges $E_t$ may be directed or undirected and weighted or unweighted. The self-edges of the type $(v_t, v_{t+1})$ are always directed.

## 2.1 Spread Blockers

We now formalize the notions of processes spreading in a network and individuals blocking this spread.

*Spread*(.) is a function that gives the overall average extent of spread in a network, that is, the expected number of individuals affected by a stochastic spreading process after a specified number of timesteps. The estimate of the spread is dependent on the model of the spreading process and the structure of the network. $Spread_v(.)$ is the expected spread in a network, when the spreading process is initiated by the individual $v$. Given a model of a spreading process $\mathcal{M}$ and a distribution of the probability of infection $\mathcal{X}$, we define the spreading functions as follows:

$$Spread_v : \{G, \mathcal{M}, \mathcal{X}\} \rightarrow \mathbf{R}^+$$

$$Spread(G) = \frac{1}{|V|} \sum_{v \in V} Spread_v(G, \mathcal{M}, \mathcal{X})$$

We define $Bl_X(.)$ as a function that measures the reduction in the expected spread size *after removing* the set $X$ of individuals from the network. Hence, the blocking capacity of a single individual $v$, $Bl_v(.)$, is the reduction in expected spread size after removing individual $v$ from the network.

$$Bl_X : G \rightarrow \mathbf{R}^+, \qquad Bl_X(G) = Spread(G) - Spread(G \setminus X).$$

$kBl(.)$ is the function that finds the set of individuals of size $k$ that results in the maximum reduction in spread in a network when that set is removed from the network. That is, this function finds the best blocker(s) in a network.

$$kBl(G) = \max_{X \subseteq V, |X| = k} Bl_X(G).$$

Thus, finding the best blockers in the network is equivalent to finding the (set of) individuals whose removal from the network *minimizes* the expected extent of spread.

$$kBl(G) = \min_{X \subseteq V, |X| = k} Spread(G \setminus X).$$

This definition of the individuals' blocking capacity by removal corresponds in the disease spread context to the quarantine action. Vaccination or inoculation leave the node in the network but deactivate its ability to propagate the spread. For the Independent Cascade model of spread (Section 2.3) the two actions are equivalent at the abstract level of estimating the spread and blockers in networks.

We next briefly define the structural network measures used in this paper.

## 2.2 Network Structural Measures

In network analysis various properties of the representing graph are studied as proxies of the properties of the individuals, their interactions, and the population. For example, the degree, various centrality measures, clustering coefficients, or the eigen values (PageRank) of the nodes have been used to determine the relative importance of the individuals [9, 26]. Betweenness centrality has been used to identify cohesive communities [18] and the distributions of shortest path lengths employed to measure the "navigability" of the network [41]. These and many other graph theoretic measures have been translated to many social properties [31, 34, 35].

The blocking capacity of an individual is one such property. Work has been done in finding efficient local strategies for vaccination and network attack [24] using clustering coefficient and assortative mixing coefficient. In this paper we evaluate the power of all the standard node properties to indicate the blocking ability of the corresponding individual. Moreover, we extend the standard static measures to reflect the dynamic nature of the underlying network. We define the following measures: degree, average degree, betweenness, closeness centralities and clustering coefficient. We modify those to incorporate the time ordering of the interactions.

We use the following terms interchangeably in this paper: individuals or nodes are the vertices of the network and interactions are edges that can be both directed or undirected. Neighbors of a node, $N(.)$, is the set of nodes adjacent to it. The subscript $T$ with a function name indicates the dynamic variant of the function.

**Degree** of a node, $DEG(.)$, is the number of neighbors of that node. We refer to the degree in the aggregate network as the aggregate degree.

**Density** is the fraction of the possible edges present in the network: $D(G) = |E|/\binom{|V|}{2}$

**Dynamic Density** average density of a timestep i.e. $D_T(G) = \sum_{1<t\leq T} D(G_t)/T$

**Diameter** is the maximum length of a shortest path.

**Temporal Path** [23] in a dynamic network is a sequence of nodes $v_1, \dots, v_p$ where each $(v_i, v_{i+1}) \in E$ is either an edge in $E_t$ for some $t$ or is a self edge $(u_t, u_{t+1})$. Also, for any $i, j$ such that $i + 1 < j$, if $v_i \in V_t$ and $v_j \in V_s$ then $t < s$. The length of a temporal path is the number of timesteps it spans.

**Dynamic Diameter** of a dynamic network is the maximum length of a shortest temporal path in the network.

**Dynamic Degree** is the change in the neighborhood of an individual over time:

$$DEG_T(u) = \sum_{1<t\leq T} \frac{|N(u_{t-1}) - N(u_t)|}{|N(u_{t-1}) \cup N(u_t)|} |N(u_t)|.$$

The dynamic degree captures the gregariousness of an individual, an important quality from a spreading perspective.

**Dynamic Average Degree** is the average over all timesteps of the interactions of an individuals in each timestep:

$$AVG\text{-}DEG_T(u) = \frac{1}{T} \sum_{1\leq t\leq T} DEG(u_t).$$

**Betweenness** of an individual is the sum of fractions of all shortest paths between all pairs of individuals that pass through this individual. It measures the importance of individuals based on their position on the shortest paths connecting pairs of non-adjacent individuals [2, 16, 17].

**Dynamic Betweenness** of an individual is the fraction of all shortest *temporal paths* that pass through it. This version incorporates the measure of a delay between interactions as well as the individual being at the right place *at the right time*. We present different flavors of the traditional betweenness centrality for dynamic networks based on position, time, and duration of interactions among individuals in [23]. The general form of the definition is as follows.

Definition 2. *Let $g_{st}$ be the number of shortest temporal paths between $s$ and $t$, $g_{st}(u)$ of which pass through $u$. Then the* temporal betweenness centrality*, $B_T(u)$, of a node $u$ is*

$$B_T(u) = \sum_{s \neq t \neq u} \frac{g_{st}(u)}{g_{st}}.$$

**Closeness** of an individual is the average distance of the individual to any other individual in the network [17, 39].

**Dynamic Closeness** of an individual is the average *time* it takes from that individual to reach any other individual in the network, based on shortest temporal paths [23].

**Clustering Coefficient** of an individual is the fraction of its neighbors who are neighbors among themselves [36].

**Dynamic Clustering Coefficient** is the sum over time of fractions of an individual's neighbors who have been interacting among themselves in previous timesteps. Let $CF(u)$ be the number of friends of $u$ that are already friends among themselves. Then the dynamic clustering coefficient

$$CC_T(u) = \sum_{0 \leq t < T} \frac{CF(u_t)}{|N(u_t)|(|N(u_t)| - 1)}.$$

**Edges in Neighborhood** is the number of edges in the local neighborhood of an individual. It loosely captures the local density of the neighborhood of an individual.

### 2.3 Spreading Model

The process of propagation in a network can be described formally using many models of transmission. For this paper we use Independent Cascade model of diffusion in networks. The independent cascade model was first introduced in [19, 20] in the context of word-of-mouth marketing. This is also the most commonly used simple model to study disease transmission in networks [11, 32, 33, 35, 37] and is closely related to the simplest Susceptible-Infectious-Recovered (SIR) models from epidemiology [1]. In the Independent Cascade model, transmission from one individual to another happens independent of interactions with all the other individuals.

The Independent Cascade model describes a spreading process comprising of two types of individuals, active and inactive. The process unfolds in discrete timesteps. In each timestep, each active individual attempts to activate each of its neighbor independently. The activation of each inactive neighbor is determined by a probability of success. If an active individual succeeds in affecting any of its neighbors, those neighbors become active in the next time step. Each attempt of activation is independent of all previous attempts as well as the attempts of any other active individual to activate a common neighbor.

More formally, the input to the independent cascade model is a set $A_0 \subseteq V$ of active individuals. In a dynamic network, an active individual $u_t$ in timestep $t$ tries to activate each of its currently inactive neighbors $v_t$ only once with a probability $p_{u_t,v_t}$, independent of all the other neighbors. If $u_t$ succeeds in activating $v_t$ at timestep $t$, then $v_t$ will be active in step $t+1$, whether or not $(u_{t+1}, v_{t+1}) \in E_{t+1}$. If $u_t$ fails in activating $v_t$, and at any subsequent timestep $u_{t+i}$ gets reconnected to $v_{t+i}$, it will again try to activate $v_{t+i}$. The process runs for a finite number of timesteps $T$. We denote by $\sigma(A_0) = A_T$ the correspondence between the initial set $A_0$ and the resulting set of active individuals $A_T$. We call the size of the set $A_T$, $|A_T|$, the *extent of spread*.

The spreading process in the independent cascade model in a dynamic network is different from the static network in one important aspect. In the static case, each individual $u$ uses all its attempts of activating each of its inactive neighbors $v$ with the same probability $p_{u,v}$ in one timestep $t$. This is the timestep right after the individual $u$ itself becomes active. After that single attempt the active individual becomes latent: that is, it is active but unable to activate others. However in the dynamic network model as defined above, the active individuals never become latent during the spreading process. For this paper, we only consider the progressive case in which an individual converts from inactive to active but never reverses (no recovery in the epidemiological model). It is a particularly important case in the context of identifying blockers since the blocking action is typically done before any recovery.

## 3. EXPERIMENTAL SETUP

We evaluate the effectiveness of each of the structural measures as indicators of individual's blocking capacity under the Independent Cascade spreading model. For each measure and for each dynamic network dataset, we follow the following steps:

1. Order the individuals according to the ranking imposed by the measure.
2. For $i = 0$ to $|V|$ do:
    (a) Remove $i$ nodes ranked top by the given measure.
    (b) Estimate the extent of spread by averaging over stochastic simulations of Independent Cascade model initiated at each node in turn, 3000 iterations for each starting node.
    (c) If the extent of spread is less than 10% of the nodes then STOP.

The probability of activation in the Independent Cascade model is set uniformly equal to the smallest probability which achieves the extent of spread of 50% of that with $Pr = 1$. The spreading process is simulated for $T$ timesteps, where $T$ is the number of timesteps in the dynamic network. The extent of spread is the number of individuals active at the end of the $T$ timesteps.

We use the following measures for comparison: dynamic and aggregate versions of degree, betweenness, closeness centralities, and clustering coefficient, as well as the average dynamic degree (turnover rate). For the datasets with directed interactions we also use page rank.Overall, we use 17 different measures.

We compare the structural measures to a random ordering of nodes as an upper bound and the best blockers as the lower bound.

## 3.1 Lower Bound: Best Blockers

We identify the best blockers one at a time using exhaustive search over all the individuals. To find one best blocker, we remove each individual, in turn, from the network and estimate the extent of spread using stochastic simulations of the Independent Cascade model in the remaining network. The best blocker, then, is the individual whose removal results in the minimum extent of spread. We then repeat the process with the remaining individuals. This process imposes another ranking on the nodes.

Optimally, one needs to identify the *set* of top $k$ blockers. However, this problem is computationally hard and an exhaustive search is infeasible. We have conducted limited experiments on the datasets considered in this paper and in all cases the set of iterative best $k$ blockers equals to the set of top $k$ blockers. This preliminary result warrants future investigation and rigorous evaluation.

# 4. DATASETS

We now describe the data sets used in our experiments.

**Grevy**: A population of Grevy's zebras (*Equus grevyi*) observed by biologists [40] during June–August 2002 in the Laikipia region of Kenya. Each node represents an individual animal and two animals are considered to be connected if their physical locations are the same.

**Onagers**: A population of onagers (*Equus hemionus*), observed by biologists [38, 40] in the Little Rann of Kutch, a desert in Gujarat, India, during January–May 2003.

**DBLP**: Co-citations among computer scientists from 1967–2005, from the *Digital Bibliography and Library Project* [30].

**Enron**: Network of e-mails sent among subpoenaed employees of the now defunct Enron corporation (available with a full description at http://www.cs.cmu.edu/∼enron/).

**Reality Mining**: Co-location of individuals in a population of MIT students; collected via mobile phone devices [14, 10].

**UMass**: Co-location of individuals in a population of students at the University of Massachusetts Amherst; data collected via portable motes(available with a full description at http://kdl.cs.umass.edu/data/msn/msn-info.html).

# 5. RESULTS AND DISCUSSION

For each of the datasets we have evaluated all the measures to determine how effectively they identify good blockers. Figure 1 shows results for two datasets, Onagers and Enron, that are representative of our results on all the datasets. The results for the other datasets are omitted due to space limitations. For all the plots, the x-axis is the number of individuals removed and the y-axis shows the corresponding extent of spread. The lower the extent of spread, the better is the blocking capacity of the individuals removed. Thus, curves lower on the plot correspond to measures that are better indicators of blocking power.

The comparison of all the measures showed that four measures performed consistently well as blocker indicators: degree in aggregate network, the number of edges in the immediate aggregate neighborhood (local density), dynamic average degree, and dynamic clustering coefficient. This is good news from the practical point of view of designing epidemic response strategies since all the measures are simple, local, and easily scalable. Figure 2 shows the results of the comparison of the four best measures, as well as the best possible and random orderings, for all the datasets. Surprisingly, while the local density and the dynamic clustering coefficients seem to be good indicators, the aggregate clustering coefficient turned out to be the worst, often performing worse than a random ordering. Betweenness and closeness performed inconsistently. Page Rank did not perform well in the only dataset with directed interactions (Enron)[1]. As seen in Figure 2, the ease of blocking the spread depends very much on the structure of the dynamic network. In the two bluetooth datasets, MIT Reality Mining and UMass, all orderings, including the random, performed similarly. Those are well connected networks. The only way to reduce the extent of spread to below 10% of the original population seems to be to trivially remove nearly 90% of the individuals. On the other hand, Enron and DBLP show the opposite trend of being easily blockable by a good ranking measure.

In addition to comparing the extent of spread based on the ranking by various measures, we compared the sets of the top ranked blockers identified by the four best measures as well as the best possible ordering. Figure 3 shows the scatter plots of the pairwise comparisons of rankings induced by the four measures. Due to space limitations, we show the two datasets where the differences are most pronounced, onagers and Enron. The scatter plots show that, in general, there is little correspondence between the rankings imposed by various measures. The only strong relationship, as expected, is between the number of edges in the neighborhood of a node and its degree in the aggregate network.

We further explore the difference in the sets of the top ranked individuals by computing the size of the common intersection of all the top sets ranked by the four measures and the best possible ranking. We use the size of the set determined by the best possible ordering as the set size for all measures. Table 1 shows the size of the common intersection for all datasets. Again, we see a strong effect of the structure of the network. The MIT Reality Mining and the UMass datasets have most of the same nodes ranked as top by all measures. On the other hand, in DBLP the four measures produced very different top ranked sets, yet all four measures were extremely good indicators of the blockers. In other networks, while there are some individuals that are clearly good blockers according to all measures, there is a significant difference among the measures. Overall, these results lead to two future directions: 1) investigating the effect of the overall network structure on the "blockability"

| Dataset | Set size | Inter. size | Inter. frac |
|---|---|---|---|
| Grevys | 5 | 2 | .40 |
| Onagres | 9 | 3 | .33 |
| DBLP | 16 | 0 | 0 |
| Enron | 13 | 4 | .31 |
| Reality Mining | 60 | 48 | .80 |
| UMass | 12 | 10 | .83 |

**Table 1: The size of the common intersection of all the top sets ranked by the four measures and the best ranking. Set size is the size of the sets determined by the best blocking ordering. The size of the intersection is the number of the individuals in the intersection and the Intersection fraction is the fraction of the intersection of the size of the set.**

---

[1] On undirected graphs, Page Rank is equivalent to degree in aggregate network

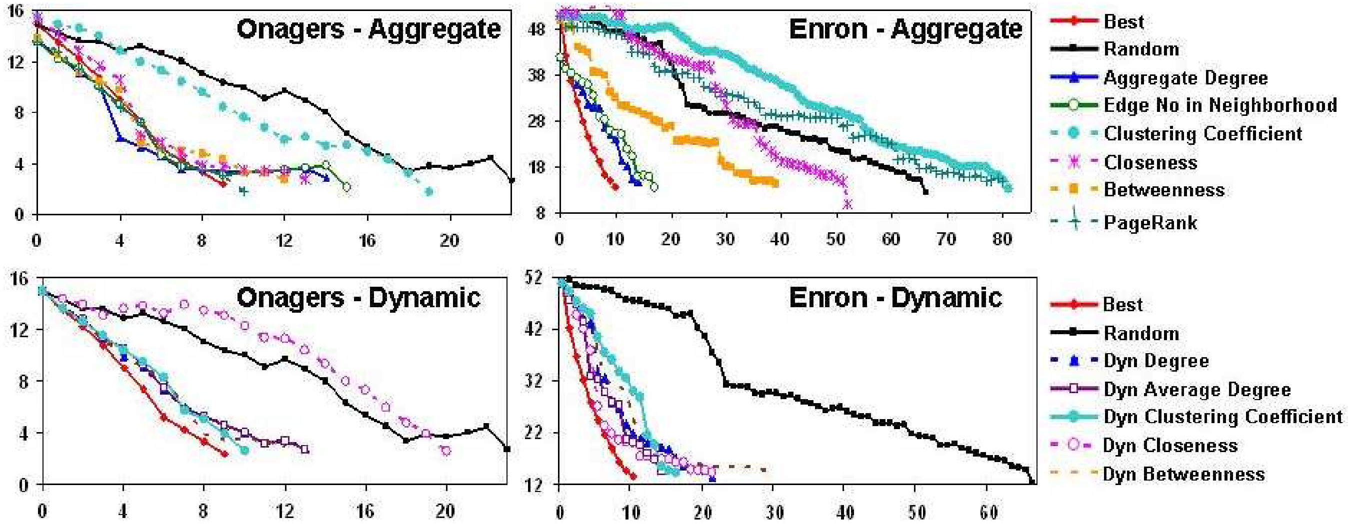

**Figure 1: [Best viewed in color.] Comparison of the reduction of extent of spread after removal of nodes ranked by various aggregate (top) and dynamic (bottom) measures in Onager (left) and Enron (right) dynamic networks. The x-axis shows the number of individuals removed and the y-axis shows the average spread size after the removal of individuals.**

of the network; and 2) designing consensus techniques that combine rankings by various measures into a possible better list of blockers.

## 6. CONCLUSIONS AND FUTURE WORK

We have studied the problem of identifying nodes in a network that are good blockers in the sense that their removal results in a large reduction in the extent of spread in a network. We empirically tested 17 different measures for identifying good blockers and determined that simple local measures, such as the degree of a node, perform as well at identifying blockers as any other measures. Moreover, these simple local measures identify good blockers on both static and dynamic networks. Several areas for future work remain, including the following. First, can we devise a theoretical model for social network formation that explains why simple local properties would so effectively predict which nodes are good blockers? Second, there are certain data sets (e.g. MIT Reality Mining and UMass datasets) in which it is inherently challenging to find good blockers, no matter what measures are used. Can we determine structural properties of the network that determine how hard it is to find good blockers?

## 7. REFERENCES

[1] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford U. Press, 1992.

[2] J. Anthonisse. The rush in a graph. *Amsterdam: Mathematische Centrum*, 1971.

[3] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *J. Comput. Syst. Sci.*, 72(6):1077–1093, 2006.

[4] A. L. Barabasi, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311(3-4):590–614, August 2002.

[5] E. Berger. Dynamic monopolies of constant size. *Journal of Combinatorial Theory Series B*, 83:191–200, 2001.

[6] K. Börner, L. Dall'Asta, W. Ke, and A. Vespignani. Studying the emerging global brain: Analyzing and visualizing the impact of co-authorship teams. *Complexity*, 10(4): 57–67, 2005.

[7] K. Börner, J. Maru, and R. Goldstone. The simultaneous evolution of author and paper networks. *PNAS*, 101(Suppl 1):5266–5273, 2004.

[8] A. Broido and K. Claffy. Internet topology: connectivity of ip graphs. In *Proceedings of SPIE ITCom*, 2001.

[9] I. Carreras, D. Miorandi, G. Canright, and K. Engøo-Monsen. Eigenvector centrality in highly partitioned mobile networks: Principles and applications. *Studies in Computational Intelligence(SCI)*, 69:123–145, 2007.

[10] A. Clauset and N. Eagle. Persistence and periodicity in a dynamic proximity network. Unpublished manuscript.

[11] Z. Dezsö and A.-L. Barabási. Halting viruses in scale-free networks. *Physical Review E*, 65(055103(R)), 2002.

[12] P. Domingos. Mining social networks for viral marketing. *IEEE Intelligent Systems*, 20:80–82, 2005.

[13] P. Domingos and M. Richardson. Mining the network value of customers. *KDD*, 2001.

[14] N. Eagle and A. Pentland. Reality mining: Sensing complex social systems. *J. of Personal and Ubiq. Computing*, 2006.

[15] S. Eubank, H. Guclu, V. Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429:180–184., 2004.

[16] L. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40:35–41, 1977.

[17] L. C. Freeman. Centrality in social networks: I. conceptual clarification. *Social Networks*, 1:215–239, 1979.

[18] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *PNAS*, 99:8271–8276, 2002.

[19] J. Goldenberg, B. Libai, and E. Muller. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 12(3):211–223, 2001.

[20] J. Goldenberg, B. Libai, and E. Muller. Using complex systems analysis to advance marketing theory development. *Academy of Marketing Science Review*, 2001.

[21] M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.

[22] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. *WWW '04*, pp. 491–501, 2004

[23] Habiba, C. Tantipanananadh, and T. Y. Berger-Wolf. Betweenness centrality in dynamic networks. Technical Report 2007-19, DIMACS, 2007.

[24] P. Holme. Efficient local strategies for vaccination and network attack. *Europhys. Lett.*, 68(6):908–914, 2004.

[25] J. Hopcroft, O. Khan, B. Kulis, and B. Selman. Natural communities in large linked networks. *KDD*, pp. 541–546, 2003.

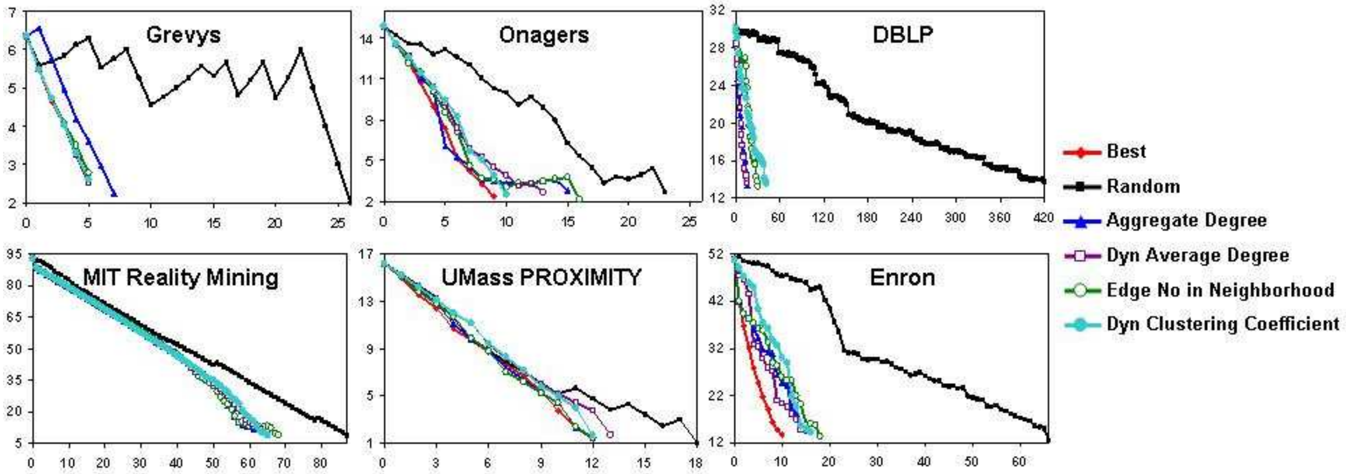

Figure 2: [Best viewed in color.] Comparison of the reduction of the extent of spread after removal of nodes ranked by the best measures. The x-axis shows the number of individuals removed and the y-axis shows the average spread size after the removal of individuals.

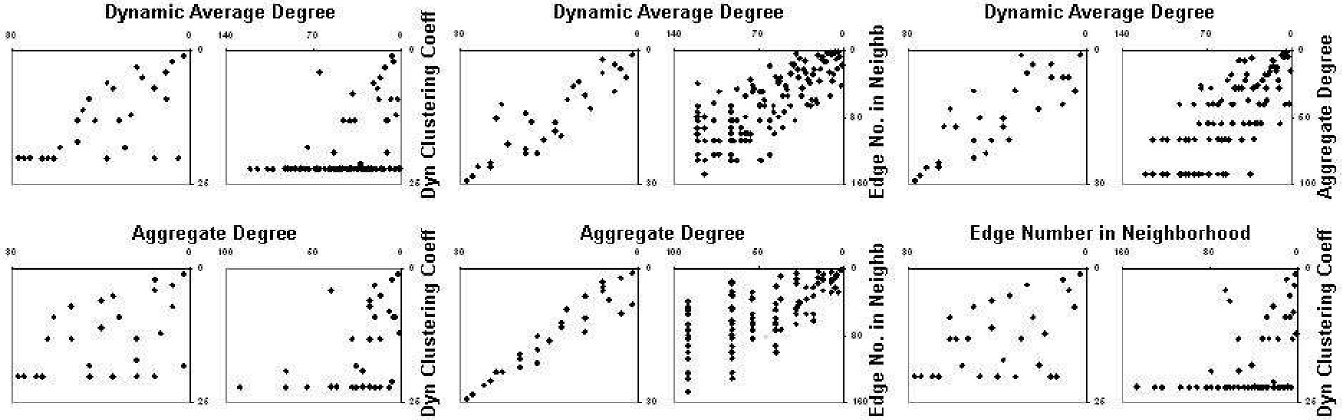

Figure 3: Comparison of the ranking of the nodes by each of the best measures. Each plot shows the ranking of the individuals according to two measures. Individuals ranked top by both measures are in the upper right corner of each plot. Nodes with the same value for a particular measure have the same rank for that measure. For each pair of plots, the left plot represents Onagers and the right plot represents Enron.

[26] F. Jordán and J. Benedek, Z.and Podani. Quantifying positional importance in food webs: A comparison of centrality indices. *Ecological Modelling*, 205:270–275, 2007.

[27] D. Kempe, J. Kleinberg, and A. Kumar. Connectivity and inference problems for temporal networks. *J. Comput. Syst. Sci.*, 64(4):820–842, 2002.

[28] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, and J. VanBriesen. Cost-effective outbreak detection in networks. *KDD*, 2007.

[29] K. Lewin. *Principles of Topological Psychology.* New York: McGraw Hill, 1936.

[30] M. Ley. Digital bibliography & library project (DBLP). `http://dblp.uni-trier.de/`, December 2005. A digital copy of the databse has been provided by the author.

[31] F. Liljeros, C. Edling, and L. N. Amaral. Sexual networks: Implication for the transmission of sexually transmitted infection. *Microbes and Infection*, 2003.

[32] R. M. May and A. L. Lloyd. Infection dynamics on scale-free networks. *Physical Review E*, 64(066112), 2001. DOI: 10.1103/PhysRevE.64.066112.

[33] M. Morris. Epidemiology and social networks:modeling structured diffusion. *Sociological Methods and Research*, 22(1):99–126, 1993.

[34] M. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[35] M. E. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66(016128), 2002. DOI: 10.1103/PhysRevE.66.016128.

[36] M. E. J. Newman. Scientific collaboration networks. i. network construction and fundamental results. *Physical Review E*, 64:016131, 2001.

[37] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86(14):3200–3203, 2001.

[38] D. I. Rubenstein, S. Sundaresan, I. Fischhoff, and D. Saltz. Social networks in wild asses: Comparing patterns and processes among populations. In A. Stubbe, P. Kaczensky, R. Samjaa, K. Wesche, and M. Stubbe, editors, *Exploration into the Biological Resources of Mongolia*, volume 10. Martin-Luther-University Halle-Wittenberg, 2007.

[39] G. Sabidussi. The centrality index of a graph. *Psychometrika*, 31:581–603, 1966.

[40] S. R. Sundaresan, I. R. Fischhoff, J. Dushoff, and D. I. Rubenstein. Network metrics reveal differences in social organization between two fission-fusion species, Grevy's zebra and onager. *Oecologia*, 2006. doi 10.1007/s00442-006-0553-6.

[41] D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.