

# SCIENTIFIC REPORTS



OPEN

## Benchmarking Measures of Network Influence

Aaron Bramson<sup>1,2,3</sup> & Benjamin Vandermarliere<sup>2,4</sup>

Received: 18 April 2016

Accepted: 05 September 2016

Published: 27 September 2016

Identifying key agents for the transmission of diseases (ideas, technology, etc.) across social networks has predominantly relied on measures of centrality on a static base network or a temporally flattened graph of agent interactions. Various measures have been proposed as the best trackers of influence, such as degree centrality, betweenness, and  $k$ -shell, depending on the structure of the connectivity. We consider SIR and SIS propagation dynamics on a temporally-extruded network of observed interactions and measure the conditional marginal spread as the change in the magnitude of the infection given the removal of each agent at each time: its temporal knockout (TKO) score. We argue that this TKO score is an effective benchmark measure for evaluating the accuracy of other, often more practical, measures of influence. We find that none of the network measures applied to the induced flat graphs are accurate predictors of network propagation influence on the systems studied; however, temporal networks and the TKO measure provide the requisite targets for the search for effective predictive measures.

In order to effectively prevent the spread of diseases one must identify those individuals with the greatest potential to change propagation outcomes. A similar need exists for efficiently spreading information across a social network. There are two main strategies to identifying the key agents for disease/idea spread: (1) the discovery of “super-spreaders”<sup>1–6</sup> and (2) finding effective immunization/removal targets<sup>7–9</sup>. The difference is not the goal of the analysis; both approaches seek to ascertain the actual or potential *influence* of each node on network propagation by performing an isolated contingency analysis. The first approach is some version of variably seeding an infection and determining how well it spreads in each setup<sup>10–12</sup>. The second approach is some version of setting nodes as firewalls and measuring changes in how the property/idea/disease spreads with different firewalls<sup>7</sup>. By toggling the status of any one node and examining the differences it generates one can ask, “How much of the propagation is this node responsible for?” Here we propose a measure called “temporal knockout” (TKO) that combines the super-spreader and immunization approaches and also includes the timing of infections to more accurately measure each agent’s influence/impact on the propagation.

The dominant technique to assess individual influence is to take a set of agents and a network of potential interactions among them and simulate the propagation of a property using a variation of SI/SIR/SIS dynamics across the network to see how far and how fast it spreads. There are variations in the (generated or empirical) network structure used, the number and placement of initial infections, the disease parameters, and with these there are variations in the identified best measure of influence (see Danon<sup>13</sup> for an extensive review on the possible variations). The most important lesson from these analyses is that different structures make different targets more effective for immunization. For example, connectivity on some network structures is resilient to random node removals but sensitive to targeted removal of nodes with certain properties, such as high degree agents in scale-free networks<sup>14–16</sup>. For other network structures, high degree is not the best measure of importance; betweenness,  $k$ -core, and other measures have been proposed as capturing key individuals in certain specific network structures and real-world datasets<sup>17</sup>. Recent work has responded to the inadequacy of traditional centrality measures by developing novel measures such as dynamical influence<sup>18</sup>, disease spreading walks<sup>19</sup>, accessibility<sup>20</sup>, epidemic centrality<sup>6</sup>, and expected force<sup>21</sup> among others; although they each share similarities to the common measures or combinations of them.

In order to evaluate network measures’ ability to track influence one must have an independent assessment of that influence – the ground truth to be matched. A common way to measure this is to seed the initial infection at each node and measure the resulting spread, typically as the cumulative cases for SIR. However, an individual’s

<sup>1</sup>Riken Brain Science Institute, Laboratory for Symbolic Cognitive Development, Wako, 351-0198, Japan. <sup>2</sup>Ghent University, Department of General Economics, Ghent, 9000, Belgium. <sup>3</sup>University of North Carolina at Charlotte, Department of Software and Information Systems, Charlotte, 28223, USA. <sup>4</sup>Ghent University, Department of Physics and Astronomy, Ghent, 9000, Belgium. Correspondence and requests for materials should be addressed to A.B. (email: bramson@brain.riken.jp)

impact on the dynamics of propagation on complex networks is more nuanced than these simple propagation measures indicate. Even when a disease starts at node  $x$ , some later-infected node  $y$  may be more responsible for the scope of the spread. In actual disease propagation dynamics<sup>22,23</sup> it is also possible that an agent being infected early *reduces* the eventual scope of the infection by altering the set of individuals that agent comes in contact with while infected.

In light of these possibilities it is clear that one must analyze how the full dynamics unfold in order to correctly assess influence over those dynamics. To incorporate the temporal aspect into our influence analysis we capture the infection propagation in a temporally extruded network structure called a “temporal web” – a variant of temporal networks<sup>24,25</sup> in which the interactions extend across time creating a single acyclic digraph rather than layered networks<sup>26–28</sup>. This temporal web provides a time-extruded version of cumulative cases that we call “magnitude” combining both the number of infected individuals and the length of their infections<sup>5</sup>.

To perform the isolated contingency analysis we propose a measure called “temporal knockout” (TKO) that combines the super-spreader and immunization approaches and also includes the timing of infections to more accurately measure each agent’s influence/impact on the propagation. TKO is not an alternative network measure for approximating influence, but rather an all-things-considered empirical measurement of each agent’s time-dependent potential to change propagation outcomes for use as a benchmark in evaluating network measures.

First we explain the temporal web construction in more detail, then we describe the process to calculate the disease magnitude and temporal knockout score. Because the temporal knockout score calculation is computationally expensive, it is desirable to have a simpler proxy measure, or set of proxy measures, that accurately reflects agent influence. Toward this end we run a battery of experiments on small world and scale-free networks and evaluate the effectiveness of some standard and newer flat/static network measures to capture influence using the TKO scores as a benchmark measure. The evaluation of network measures presented here is indicative of the need for improved ways to capture propagation influence, but our focus here is the presentation of TKO as a standardized benchmark metric for performing such investigations.

## Approach

Our analysis proceeds through the following steps: (1) create collections of scale-free and small world base networks; (2) build temporal webs encapsulating a fixed set of potential interactions for each one; (3) simulate propagation dynamics across each temporal web for each agent of each network; (4) calculate the temporal knockout of each node in the temporal web; (5) generate the flattened network and analyze the flat networks using centrality measures; (6) examine the degree to which the flat network measures capture the agents influence as measured by TKO.

**Network and Disease Parameters.** We simulate the spread of an infectious disease using an agent-based model realizing SIR and SIS dynamics. Our networks have 200 agents connected in either a small world or scale free network with 800 and 784 edges respectively. The small world base networks are undirected connected Watts-Strogatz networks where each agent is connected to  $k = 8$  neighbors and the probability of rewiring is set to  $p = 0.025$ . The scale-free base networks are undirected Barabasi-Albert networks with  $m = 4$  as the number of edges to attach from a new node to existing ones. The networks were generated using the implementation of the python package NetworkX<sup>29</sup>. For each combination of network type and infection probability ( $p = 0.10, 0.15$ , and  $0.20$ ), we generate 25 instantiations (150 total). We note that the SIR and SIS versions of a given combination run on the same instantiations, thus using the same link activations at each time step. In each iteration of the model, the probability that a given link is activated is

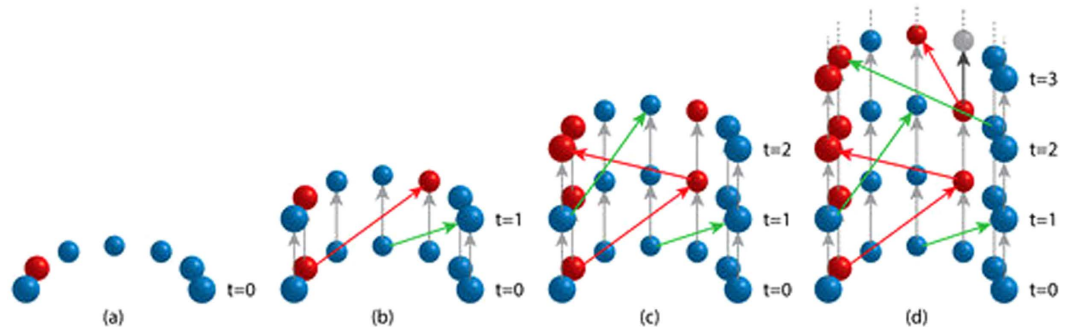
$$p_{ij} = \frac{k_j^{-1}}{\sum_{K_i} k_n^{-1}}$$

with  $k_j$  being the undirected degree of agent  $j$ , and the summation in the denominator is over each network neighbor ( $n$ ) of node  $i$  (written  $K_i$ )<sup>30</sup>. There is one initially-infected agent per run and we perform a run of the model using each agent as the initial agent for each of the 25 instantiations of each network type. Each infectious agent has a probability to infect susceptible network neighbors and we run the full battery of simulations using infection probabilities of 0.10, 0.15, and 0.20. In each period, the probability of infectious agents converting to recovered/susceptible ( $I \rightarrow R$  and  $I \rightarrow S$  for SIR and SIS models respectively) is  $1/15$ . Each run lasts 200 periods; this is typically sufficient for SIR dynamics to run their course, and is used for SIS models for parsimony of analysis.

**Building a Temporal Web.** We run our simulations using simultaneous updating so that each agents’ state at  $t + 1$  depends on their state at  $t$  and interactions initiated at  $t$ . When represented as an intertemporal network the interaction edges therefore run across time from agents at  $t$  to other agents at  $t + 1$  in addition to “inheritance edges” from each agent at  $t$  to its  $t + 1$  self (see Fig. 1). We call this version of intertemporal networks a “temporal web” because it produces a single acyclic directed graph across time rather than connected layers.

We first build the temporal web “skeleton” that includes all of the state changing and interaction probabilities which may be needed for any particular run. With non-adaptive interaction probabilities, who interacts with whom *and when* all become fixed for those runs. Therefore when we run the simulation using each agent as the initially infected agent, the overall dynamics are kept constant while we monitor the propagation so that the only difference is the initial agent.

**Disease Magnitude.** The temporal structure facilitates a variety of new measures, which are defined and explored elsewhere<sup>26,31,32</sup>. Specifically for epidemiology it becomes natural to switch to a temporally extended



**Figure 1.** A simplified example of building a “temporal web” style intertemporal network from state-change and interaction data for an SIR model. This procedure differs from temporally layered networks in that the interaction edges are cross-temporal to capture simultaneous updating in the generated data, thus creating a single acyclic directed graph across time.

refinement of the standard cumulative cases measure. Rather than (or in addition to) reporting the number of agents that are ever infected, the disease *magnitude* is calculated as the number of agent-times (i.e., nodes in the temporal web) that are in the infectious (or exposed) state. It is equivalent to the cumulative sum of the number of infectious agents across iterations<sup>5</sup>. This measure better captures disease morbidity because it accounts for both the number of infections and how long the infections persist – a large number of very short infections could be considered preferable to a few persistent long-term infections. Depending on the application, the node count or a normalized version may be preferable – the number of nodes is the same for all of our experiments described below, so we use the “raw magnitude.”

**Calculating the Temporal Knockout Scores.** Temporal knockout (TKO) measures influence by aggregating two levels of contingency. First we select an agent from the population to be initially infected and run the disease model while capturing each agent’s state and interactions at each iteration in a temporal web. The resulting collection of infectious nodes (agent-times) embodies the magnitude of the illness contingent on that agent being the initially infected one. Then the second layer is that for each infected node in the temporal web generated by that run we perform a knockout analysis: remove that node and run *the same infection dynamics* and measure the difference in the disease magnitude. Thus for each node we capture the change in disease magnitude contingent upon that agent being removed at that time, contingent upon that particular initially infected agent.

The initially infected agent at the  $t_0$  iteration will have a marginal infection effect equaling the whole magnitude. Note that removing a noninfectious later at  $t_0$  still prevents it from being infected later, which affects the marginal infection score of that agent at  $t_0$ ; however, the pre-infection time nodes for an agent will have the same TKO as the first infected time node; thus the calculation can be performed on just the infected subset and backtracked to earlier times. Perhaps counter-intuitively this effect can be negative; i.e., it is possible to remove an agent from the system at a particular time and have the overall disease spread *increase*. This can happen when agents that are infected by the knocked out agent would normally have quickly lead to dead ends, but when instead infected later by other agents they spread the disease to many more others.

We perform this knockout analysis for every node in the temporal web to get the marginal infection score conditional on that initial agent. We repeat this process using each of the agents as the initially infected agent and set each node’s TKO score as the average marginal infection score across those runs. Thus we have the conditional marginal infection spread for each agent at each time step for all possible single-agent disease carrier initial conditions. This algorithm therefore captures the *potential* for each agent at each period to influence the spread of the disease.

Because TKO is an overt counting of infected agent-times given the contingent hypothetical-empirical results instead of a summary measure we believe that it stands as a reliable benchmark for the influence of each agent (in networked epidemiological systems). Also note that TKO’s hypothetical-empirical approach means that the change in total infection after a knockout of agent  $A_i$  at any time  $t_r$  cannot be calculated except through the resimulation of the infection dynamics across the rest of the temporal web. Because of this TKO is thoroughly descriptive of the observed dynamics, but it is not predictive of influence in other runs.

**Base and Flattened Graphs.** In order to predict which agents are most likely to facilitate diffusion, we wish to compare the TKO identification with measures on flat, non-temporal networks. Specifically we would like to know how well each of various centrality measures does in capturing each agent’s network influence as benchmarked by TKO. Two versions of flat graphs are relevant here: (1) the base potential interaction network from which the actual interactions were probabilistically generated and (2) the flattened empirically observed interactions. Our results for the base network and weighted and unweighted flattened networks are nearly identical, so we focus on the base network here and leave the flattened networks for the Supplementary Materials. We have twenty-five distinct base networks for each scenario (although each SIR and SIS pair use the same networks) and for every node in each one we calculate the following centrality and influence measures:  $k$ -core, degree, closeness, betweenness, eigenvector, and Katz centralities, accessibility (path length = 2), and expected force.

Infection Type	Network Type	Infection Probability	Mean Magnitude	Magnitude StDev	Percent Duds
SIR	Scale Free	0.10	143.352	288.549	0.625
SIR	Scale Free	0.15	584.744	774.628	0.482
SIR	Scale Free	0.20	1296.44	1142.24	0.380
SIR	Small World	0.10	88.9266	131.743	0.584
SIR	Small World	0.15	227.321	324.207	0.457
SIR	Small World	0.20	445.033	559.017	0.352
SIS	Scale Free	0.10	548.746	1155.19	0.593
SIS	Scale Free	0.15	5003.03	5237.44	0.445
SIS	Scale Free	0.20	10800.6	8150.76	0.344
SIS	Small World	0.10	308.734	536.106	0.557
SIS	Small World	0.15	2526.97	2839.26	0.433
SIS	Small World	0.20	7036.79	5623.18	0.333

**Table 1. Results summary of infection spread for each model variation.** Each row aggregates 5000 runs (one run initialized at each of 200 agents for each of the 25 base network implementations). Duds are defined as runs in which the raw magnitude is fewer than 50 agent-times.

## Results

The infection dynamics in our model match other models with similar network structures and disease parameters<sup>13,30</sup>. We briefly summarize the contagion results in order to provide context for the centrality measures and to facilitate comparisons to other models. For our SIR models the cumulative cases and magnitude measures are nearly perfectly correlated (0.995) because the fixed 1/15 probability of I → R transitions implies a uniform *expected/average* infection duration time of 15 iterations. For SIS models reinfection can multiply an agent's contribution to magnitude but still only be counted once by the number of cumulative cases, so the correlation is reduced (0.936), but is still high due to the relatively short time horizon for our SIS simulations (200-iterations).

As seen in Table 1 both network types show high variation in magnitude depending on the initial agent; however, when aggregated across the 25 implementations of each network type they reveal similar magnitude profiles (see Supplementary Materials for details). For ease of reading we present the raw (non-normalized) magnitude scores (i.e., the number of infectious nodes in the temporal web). As you can see in Table 1 there are a large number of runs in which the disease never catches on (what we call “duds”) and although these outcomes drag the mean magnitude down and raise the variance, for our purposes there is no benefit in separating out the duds and, for example, testing the remaining infections for matches to known distributions because we do not utilize these summary statistics in any of our TKO analyses.

We also calculate the “epidemic probability” for each agent as a binary variable for whether the run reaches a magnitude greater than 50 when starting at that node. Because, unlike cumulative cases, the magnitude can vary greatly between runs with identical outcomes (e.g., full saturation) due to timing effects the correlation between the magnitude values and epidemic probability is only 0.617. Furthermore, the high variance in magnitude scores (even just among non-dud runs) is large enough to undercut the reasoning for preferring epidemic probability over a fully quantitative measure such as magnitude in this case<sup>21</sup>.

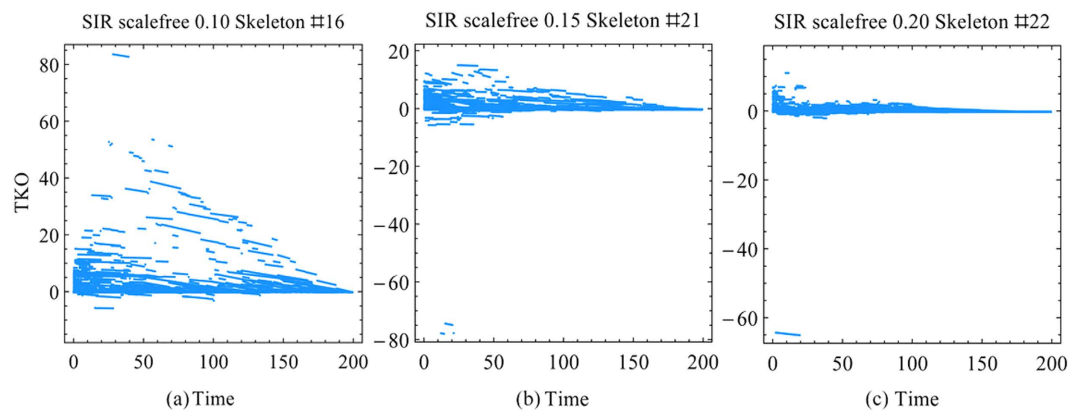
**TKO vs Magnitude Correlations Results.** We first compare the TKO score of each agent to the initial-agent resulting magnitude in order to evaluate whether this standard measure of influence effectively captures a node's ability to spread disease. The TKO algorithm accounts for the idiosyncrasies of the agent interactions across time, but as a result it assigns scores across time as well. In order to compare TKO node scores to initial-agent-spread scores we first need to aggregate them to the individual agents.

For each node we determine two versions of TKO: (1) the proportional change in the number of infectious nodes and (2) the change in the fraction of nodes that become infectious. The proportional change of node  $i$  is calculated as the number of agents that are infected when node  $i$  has been removed divided by the number of nodes that were originally infected, and then that subtracted from one so that a value of one means that no nodes become infected if this one is removed. Alternatively the delta fraction is the fraction of infected nodes in the original run minus the fraction of nodes that become infected when node  $i$  is removed. For both versions negative values occur when more nodes become infected contingent upon  $i$ 's removal compared to the original run. An agent that was never infected will have a TKO value of zero for all its temporal nodes. For each of these temporal node-based measures we aggregate them to agents by considering both the maximum value an agent achieves across time and its average TKO score across time.

The Pearson correlations for agent TKO scores and magnitude appear in Table 2. In the most correlated scenario (SIR smallworld 0.10 infection rate) the best match is to maximum TKO with a correlation coefficient just under 0.50 (marked with\*). Although we initially believed that the Spearman rank correlations would be higher, they are actually very similar and not consistently better or worse (a table of Spearman correlations appears in the Supplementary Materials). For example, the best-case scenario for the Spearman correlation is the same, with a Spearman rho value of 0.517. For both types of correlation the performance drops dramatically as the disease magnitude increases (via higher infection rates), indicating that the large proportion of runs with almost no spread (“duds”) are trivially improving the correlations and overstating the ability of agent-initiated magnitude to measure propagation impact.

Disease Type	Network Type	InfectionRate	MaxProportion	MaxDeltaFraction	AveProportion	AveDeltaFraction
SIR	scalefree	0.10	0.403	0.405	0.288	0.292
SIR	scalefree	0.15	0.067	0.246	0.064	0.157
SIR	scalefree	0.20	0.046	0.219	0.078	0.158
SIR	smallworld	0.10	0.494*	0.472	0.366	0.364
SIR	smallworld	0.15	0.043	0.265	0.077	0.189
SIR	smallworld	0.20	0.03	0.192	0.015	0.119
SIS	scalefree	0.10	0.347	0.376	0.268	0.282
SIS	scalefree	0.15	0.057	0.248	0.084	0.153
SIS	scalefree	0.20	0.045	0.234	0.059	0.108
SIS	smallworld	0.10	0.404	0.418	0.353	0.371
SIS	smallworld	0.15	0.024	0.201	0.045	0.15
SIS	smallworld	0.20	0.05	0.19	0.042	0.107

**Table 2.** The mean Pearson correlation coefficients across the 25 network instantiations of the disease magnitude given an agent is the initially infected agent and the TKO scores for that agent. The low correlations imply that using the disease spread based on initial infection is a poor measure of influence.



**Figure 2.** Plot of TKO scores across time for SIR dynamics and a scalefree network. These examples show that the most influential agent-times often do not occur during the initial phases of a disease, but can indicate bottlenecks in the spread of the disease. This also shows the appearance of negative TKO agents, the removal of which actually *increases* the morbidity of the disease due to timing and network effects.

We also compare agent TKO to epidemic probability and both the Pearson and Spearman correlations of this analysis are very similar overall to the comparison to magnitude, with a maximum Pearson correlation of 0.323 and a maximum Spearman correlation of 0.518 (full tables appear in the Supplementary Materials). Overall they are a weaker match with TKO, but this should be expected because TKO itself is calculated from aggregated contingent marginal magnitudes.

The poor correlations between TKO and both agent-initiated magnitude and epidemic probability have multiple explanations. To understand the relationship better we present a few select plots of the agent TKO scores across time in Fig. 2. These plots present the change in magnitude resulting from removing each infectious agent at each time averaged across the 200 runs initialized with each agent being infected. So a value of  $m$  means that *on average* (i.e., regardless of which agent is initially infected) removing this agent at this time decreases morbidity by  $m$  agent-times.

As we saw, there are many dud runs in which the disease doesn't spread beyond a few initial agents; such cases bring down the average values but they remain comparable across different infection scales here because all our simulations have the same number of runs. A TKO score of twenty might mean 500 saved agent-times in one run and none in the others, or 50 in ten runs, etc. So TKO scores can be small if the disease tends not to spread much because no agent at no time will be a key player in the localized infections. On the other hand, if the disease spreads rapidly from every agent to the whole population then no single agent could be particularly responsible for the scale of the infection across multiple initializations. So TKO will be small in these cases too because there are just too many infection paths for any one agent to be a key player on enough of them to have a high knockout effect. Thus, unlike measuring influence via cumulative cases (or magnitude or epidemic probability) in which every agent may be seen as influential, TKO scores are high only if an agent is influential in the sense of playing a key role in the amount of spread. This difference explains why correlations drop as the infection size increases.

Disease Type	Network Type	InfectionRate	Degree	Closeness	Betweenness	Eigenvector	Katz	Accessibility	Expected Force
SIR	scalefree	0.10	0.127	0.106	0.099	0.098	0.08	0.081	0.119
SIR	scalefree	0.15	0.125	0.096	0.102	0.098	0.084	0.069	0.114
SIR	scalefree	0.20	0.14	0.119	0.115	0.115	0.101	0.096	0.133
SIR	smallworld	0.10	-0.005	0.031	0.019	-0.008	-0.015	0.039	-0.006
SIR	smallworld	0.15	0.049	0.07	0.082	0.08	0.064	0.091	0.051
SIR	smallworld	0.20	0.067	0.127	0.102	0.026	0.057	0.136	0.062
SIS	scalefree	0.10	0.135	0.089	0.111	0.091	0.068	0.061	0.111
SIS	scalefree	0.15	0.162	0.098	0.127	0.106	0.085	0.054	0.13
SIS	scalefree	0.20	0.233	0.167	0.191	0.168	0.138	0.107	0.201
SIS	smallworld	0.10	0.025	0.054	0.02	-0.013	0.004	0.041	0.022
SIS	smallworld	0.15	0.068	0.102	0.11	0.025	0.056	0.107	0.066
SIS	smallworld	0.20	0.123	0.174	0.241	0.052	0.126	0.232	0.121

**Table 3. The Pearson correlations between the mean proportional TKO score with each of five base network agent centrality scores.** Tables for the other results appear in the Supplementary Materials.

Up to this point we have argued that temporal magnitude is a more accurate measure of disease morbidity than cumulative cases because magnitude accounts for both the length of infection as well as agent reinfections. A network measure's ability to capture an agent's influence on disease is standardly compared to the eventual spread of the disease contingent upon it starting at that agent, but our analysis of correlations with TKO shows that this standard measure of impact itself fails to capture how much disease spread that agent is responsible for because it lacks sensitivity to the structure of the interactions across time. Specifically, agent-initiated metrics cannot account for the role that some other agent later has in the spread of disease, nor how consistently important a particular agent is for spread regardless of the initial infection. From these results we tentatively conclude that TKO stands as the best measure of an agent's influence on network propagation. We now turn to testing the ability of static network measures to identify a system's high-impact agents.

**Predicting Temporal Knockout from the Static Interaction Network.** The temporally extruded network structure captures the system dynamics in a way that facilitates contingency analyses, however one must already have the data across time to measure those properties, including TKO. For predictive purposes we would like to know if there is some property of the known interaction structure that can identify key players<sup>8</sup>. Although temporal networks are gaining popularity (see Holme<sup>25</sup> for a review), most network analysis is still performed on flat networks because there are already measures available with known interpretations. The question here is whether any flat graph property can accurately predict the conditional marginal infection as measured by agent-aggregated temporal knockout.

We ran the three comparisons between each of the four aggregated TKO measures and each of eight network centrality measures. Both Pearson and Spearman correlations were calculated. Furthermore, because the standard network centrality measures only purport to capture the highest value agents properly (i.e., rather than a claim to assigning accurate values to all nodes) we also compared the overlap between the ten agents (5%) with the top TKO scores with the ten agents with the top centrality scores<sup>17</sup>. We compared the maximum proportional and maximum delta fraction TKO as well as the average proportional and average delta fraction TKOs with degree, closeness, betweenness, eigenvector, and Katz centrality, as well as accessibility and expected force ( $k$ -core values were too undifferentiated on our base networks to be meaningful and are omitted here). The full output of the analysis appears in the Supplementary Materials, but they are qualitatively similar enough that Table 3 suffices to understand the general results.

We find that neither the Pearson nor the Spearman correlations are systematically higher, nor is any one of the network measures consistently better than all the others (although eigenvector and Katz centrality are typically worse). Notably, accessibility and expected force, two newer measures specifically designed to measure epidemiological spread, do not fare better than the common centrality measures. Although the correlations are typically positive, the correlation coefficients and Spearman Rhos are almost entirely below 0.20 and there are zero instances across all results of relations above 0.4. Differences between the proportional and delta fraction TKOs are small (as expected), but not negligible; delta fractional correlations tend to be better but not in every case. Similarly, the correlations with mean TKO tend to be slightly higher than maximum TKO, but the differences are small and inconsistent. For the top ten overlap comparison we find that the centrality measures typically find a few of the top ten TKO agents, with the highest average matching score of 0.212.

There are other patterns in the results that may offer clues to where to look for improved network measures. For example, for each disease type, each network type, and each TKO version the correlations of all measures tend to be higher with larger infection rates. Unsurprisingly, degree centrality typically performs better on the scalefree networks than the small world networks. However, any such pattern may be spurious because the correlation values are too low and similar for our sample size to provide adequate power. In summary, the result is that none of the eight measures we consider on the flat interaction network can predict which agents have the greatest influence on spreading a disease.

## Conclusions

In this paper we have argued that using temporal networks to capture disease spread has the benefits of incorporating the details of the interaction timing which is necessary for judging each agent's level of influence/impact on the spread. The number of infectious agent-time nodes, a measure we call magnitude, is superior to cumulative cases because it captures both the length of infections and agent reinfection. However, adapting the standard measures of influence – eventual spread contingent upon the starting agent or blocked spread contingent upon removing the agent – to magnitude is insufficient to properly capture an agent's overall level of influence. Although eliminating the initial agent is a sure-fire way to stop the spread, that is not informative for deciding whom to remove before the disease starts. What is needed is the change in the spread of disease contingent upon each agent being removed generalized over all possible initial agents. But the degree of influence is also dependent on *when* the agent is removed because the interaction dynamics of these systems are complex: removing an agent early can increase the eventual spread. We present the temporal knockout measure to capture all these contingencies and provide a general benchmark for propagation influence.

One key insight from this study is that an agent's influence depends on how the dynamics unfold through time, which cannot be accurately predicted by historic interaction data or known communication channels. Nascent measures on temporal network structure (i.e., ones that operate on the full temporal web) can accurately track the TKO property with considerably less computational time, but they still require knowing the complete interaction structure over time<sup>26</sup>. Thus, they work as effective *proxy* measures of TKO on existing temporal webs, but are not viable *predictor* measures of TKO from base graphs. Although we do not have improved static network measures to offer at this stage, we believe that having a proper benchmark for such measures provides the foundation necessary for developing them.

For most realistic health applications, by the time an intervention occurs there are already several infectious individuals, and for this reason there is interest in measures/strategies for scenarios with multiple initially infected agents<sup>13</sup>. The problem is in the combinatorics; e.g., instead of 200 runs per network, with two initial agents it becomes  $\binom{200}{2} = 19,900$  runs – for just three initial agents it becomes 1,313,400 runs. Because TKO generalizes marginal conditional spread of every agent-time across all initially infected agents, the TKOs scores can be combined post hoc without needing to rerun the simulations. So, although the TKO algorithm is computationally intense compared to the single initial agent runs, there would be considerable time savings when compared to testing every combination of initially infected agents.

As noted by Kitsak<sup>17</sup>, when using cumulative cases to capture the influence of particular agents it makes sense to keep the infection probabilities small enough that the disease typically will not spread to the whole population – otherwise the role of any single individual will be difficult to discern. TKO does not suffer from this limitation because the disease magnitude measure also detects delays in infection even if the whole population does eventually get infected. Again, the timing of the interactions is important, so in addition to facilitating a reduction in morbidity, TKO is useful for developing adaptive intervention strategies.

Recent papers have introduced new measures with claims of increased accuracy (at least in certain contexts). However, those accuracy claims are based on how well their own measure matched their own chosen metric on their own chosen network and spread parameters. We propose that TKO, in its exhaustive marginal contingent effect calculation, can act as a benchmark metric against which the performance of proposed measures can be judged – essentially establishing a ground truth for the influence of each agent (at each time) in a network.

We acknowledge that the version of temporal knockout presented here is not the only option for benchmarking epidemiological network studies. One direction of refinement is to develop measures of TKO based on thresholds of infection size changes instead of magnitude – a similar move to using epidemic probability instead of agent-initiated cumulative case<sup>21,33</sup>. Another direction is to expand the breadth of the simulations to more closely approach an exhaustive analysis of interaction possibilities, perhaps including a notion of maintaining high TKO through variations in the infection rate and disease variation into the the measure of influence<sup>6</sup>. We visit these ideas in follow-up research to establish shared benchmarks for evaluating measures of network influence on a variety of standardized generated and empirical networks similar to how Zachary's Karate Club has been used to test community detection methods. Before such benchmark networks can be established, we as a community must agree on what counts as a measure of influence. We propose that temporal knockout may fill that role, and at the very least is a useful step in the right direction.

## References

1. Kempe, D., Kleinberg, J. & Tardos, É. Influential nodes in a diffusion model for social networks. In *Automata, languages and programming*, 1127–1138 (Springer, 2005).
2. Wang, Y., Cong, G., Song, G. & Xie, K. Community-based greedy algorithm for mining top-k influential nodes in mobile social networks. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1039–1048 (ACM, 2010).
3. Kimura, M., Saito, K., Nakano, R. & Motoda, H. Extracting influential nodes on a social network for information diffusion. *Data Mining and Knowledge Discovery* **20**, 70–97 (2010).
4. Chen, D., Lü, L., Shang, M.-S., Zhang, Y.-C. & Zhou, T. Identifying influential nodes in complex networks. *Physica a: Statistical mechanics and its applications* **391**, 1777–1787 (2012).
5. Saito, K., Kimura, M., Ohara, K. & Motoda, H. Efficient discovery of influential nodes for sis models in social networks. *Knowledge and information systems* **30**, 613–635 (2012).
6. Sikic, M., Lancic, A., Antulov-Fantulin, N. & Stefancic, H. Epidemic centrality – is there an underestimated epidemic impact of network peripheral nodes? *The European Physical Journal B* **86**, 1–13 (2013).
7. Chen, Y., Paul, G., Havlin, S., Liljeros, F. & Stanley, H. E. Finding a better immunization strategy. *Physical review letters* **101**, 058701 (2008).
8. Yu, Y., Berger-Wolf, T. Y., Saia, J. et al. Finding spread blockers in dynamic networks. In *Advances in Social Network Mining and Analysis*, 55–76 (Springer, 2010).

9. Kuhlman, C. J., Kumar, V. A., Marathe, M. V., Ravi, S. & Rosenkrantz, D. J. Finding critical nodes for inhibiting diffusion of complex contagions in social networks. In *Machine Learning and Knowledge Discovery in Databases*, 111–127 (Springer, 2010).
10. Newman, M. E. Spread of epidemic disease on networks. *Physical review E* **66**, 016128 (2002).
11. Newman, M. E., Forrest, S. & Balthrop, J. Email networks and the spread of computer viruses. *Physical Review E* **66**, 035101 (2002).
12. Dekker, A. H. Network centrality and super-spreaders in infectious disease epidemiology. In *20th International Congress on Modelling and Simulation (MODSIM2013)* (2013).
13. Danon, L. *et al.* Networks and the epidemiology of infectious disease. *Interdisciplinary perspectives on infectious diseases* 2011 (2011).
14. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000).
15. Callaway, D. S., Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters* **85**, 25, 5468–5471 (2000).
16. Pastor-Satorras, R. & Vespignani, A. Immunization of complex networks. *Physical Review E* **65**, 036104 (2002).
17. Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nature Physics* **6**, 888–893 (2010).
18. Klemm, K., Serrano, M., Eguluz, V. M. & Miguel, M. S. A measure of individual role in collective dynamics. *arXiv preprint arXiv:1002.4042* (2010).
19. Bauer, F. & Lizier, J. T. Identifying influential spreaders and efficiently estimating infection numbers in epidemic models: A walk counting approach. *EPL (Europhysics Letters)* **99**, 68007 (2012).
20. Viana, M. P., Batista, J. a. L. & Costa, L. d. F. Effective number of accessed nodes in complex networks. *Physical Review E* **85**, 036105 (2012).
21. Lawyer, G. Understanding the influence of all nodes in a network. *Scientific reports* **5** (2015).
22. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15124–15129 (2004).
23. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013).
24. Holme, P. & Saramäki, J. Temporal networks. *Physics Reports* **519**, 97–125 (2012).
25. Holme, P. Modern temporal network theory: a colloquium. *The European Physical Journal B* **88**, 1–30 (2015).
26. Bramson, A. & Vandermarliere, B. Dynamical properties of interaction data. *Journal of Complex Networks* cnp009 (2015).
27. Michail, O. An introduction to temporal graphs: An algorithmic perspective. In *Algorithms, Probability, Networks, and Games*, 308–343 (Springer, 2015).
28. Speidel, L., Takaguchi, T. & Masuda, N. Community detection in directed acyclic graphs. *The European Physical Journal B* **88**, 1–10 (2015).
29. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, 11–15 (Pasadena, CA USA 2008).
30. Rahmandad, H. & Sterman, J. Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science* **54**, 998–1014 (2008).
31. Nicosia, V. *et al.* Graph metrics for temporal networks. *arXiv:1306.0493 [physics.soc-ph]* **1306** (2013).
32. Pfitzner, R., Scholtes, I., Garas, A., Tessone, C. J. & Schweitzer, F. Betweenness preference: Quantifying correlations in the topological dynamics of temporal networks. *Physical Review Letters* **110**, 198701 1–5 (2013).
33. Lancic, A., Antulov-Fantulin, N., Sikic, M. & Stefancic, H. Phase diagram of epidemic spreading – unimodal vs. bimodal probability distributions. *Physica A: Statistical Mechanics and its Applications* **390**, 65–76 (2011).

## Acknowledgements

We would like to thank Koen Schoors for making this collaboration possible. This work is supported by the Research Foundation Flanders (FWO-Flanders).

## Author Contributions

A.B. conceived the measures and wrote the bulk of the text, B.V. conducted and described the experiments, A.B. and B.V. both analysed the results and reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Bramson, A. and Vandermarliere, B. Benchmarking Measures of Network Influence. *Sci. Rep.* **6**, 34052; doi: 10.1038/srep34052 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016