Competing Memes Propagation on Networks: A Network Science Perspective

Xuetao Wei, Nicholas C. Valler, B. Aditya Prakash, Iulian Neamtiu, Michalis Faloutsos, and Christos Faloutsos

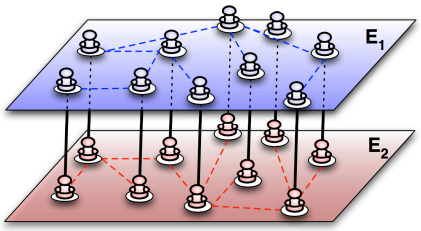
***Abstract*— In this paper, we study the intertwined propagation of two competing "memes" (or data, rumors, etc.) in a composite network. Within the constraints of this scenario, we ask two key questions: (a) which meme will prevail? and (b) can one influence the outcome of the propagations? Our model is underpinned by two key concepts, a structural graph model (composite network) and a viral propagation model (SI_1I_2S). Using this framework, we formulate a non-linear dynamic system and perform an eigenvalue analysis to identify the tipping point of the epidemic behavior. Based on insights gained from this analysis, we demonstrate an effective and accurate prediction method to determine viral dominance, which we call the EigenPredictor. Next, using a combination of synthetic and real composite networks, we evaluate the effectiveness of various viral suppression techniques by either a) concurrently suppressing both memes or b) unilaterally suppressing a single meme while leaving the other relatively unaffected.**

***Index Terms*—Epidemics, Competition, Prediction, Suppression**

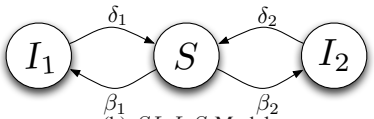
I. Introduction

IN THIS PAPER, we examine the competition of two opposed memes across interconnected agents by extend the popular *susceptible-infected-susceptible* (SIS) compartmental model to construct a novel propagation scheme. We are inspired by the popularity of epidemic models spanning various disciplines. In fact, epidemic models already accurately describe various network spreading phenomena such as the spread of social information, computer viruses, fashion trends, religious beliefs, market penetration and product adoption [1], [2], [3], [4], [5], [6]. Throughout this work, we use the general term "meme" to represent the propagating datum, but, without loss of generality, we may easily substitute the term "computer virus," "flu," or "rumor."

In our scenario, we consider two memes spreading across a population of individuals. A meme's ability to jump from individual to individual is governed by a number of factors unique to the meme itself. Thus, at a network level, the world of each meme is unique. We capture the unique network views of each competing meme through a novel graph structure we refer to as a *composite network*, illustrated in Figure 1(a). Each meme propagates across a unique plane representing a unique connectivity between individuals in our system. A composite network C is a tuple of nodes N and two edge sets E_1, E_2, i.e, $C = (N, E_1, E_2)$, where $E_1 \not\equiv E_2$. We further assume that each individual may "possess" a single meme at a time, a constraint we refer to as mutual exclusivity.



(a) Example Composite Network



(b) SI_1I_2S Model

Fig. 1. (a) Example Composite Network topology: a single set of nodes N with two distinct edge sets E_1 and E_2. (b) The SI_1I_2S State Transition Diagram, where S represents the susceptible state and $I_{\{1,2\}}$ indicate the infected state for memes M_1 and M_2. The transitions between states are indicated by the directed edges labeled $\beta_{\{1,2\}}$ and $\delta_{\{1,2\}}$.

The following example can make this problem more concrete. Consider the 2011 Egyptian revolution, which according to reports was partly coordinated via Twitter [7]. To counter such a Twitter campaign, a tech-savvy government could inject confusing and competing information using a malicious Facebook application that spreads to someone's friends. Then, which propagation will win? Hence a focal point of our paper is to predict the winner, by looking at the connectivity and the propagation behavior of its "memes." Naturally, another important question is what means one could use to influence the outcome of these competing information dissemination campaigns.

Interestingly, previous work has not focused on this problem. In fact, most previous efforts either studied a single epidemic on a single topology [8], [9], [10], [11], [12], [13], [3], or studied two pathogens, but on the same topology, and under the assumption that the two viruses appear one after the other [14]. More recent work has studied the quarantining techniques in the case of computer viruses [15], but it has not attempted to study the problem analytically or derive the

Manuscript received August 16, 2012; revised January 30, 2013.

X. Wei and I. Neamtiu are with UC Riverside (e-mail: {xwei,neamtiu}@cs.ucr.edu).

N. C. Valler is with Crowdcompass, Inc. (e-mail: nvaller@ crowdcompass.com).

B. A. Prakash is with Virginia Tech (e-mail: badityap@cs.vt.edu).

M. Faloutsos is with the Universitiy of New Mexico (e-mail: michalis@cs.unm.edu).

C. Faloutsos is with Carnegie Mellon University (e-mail: christos@cs.cmu.edu).

Digital Object Identifier 10.1109/JSAC.2013.130607.

conditions that predict the outcome in the composite networks. We discuss previous work in more detail in §VII.

In our earlier work [16], we provided a rigorous formulation of competing memes on composite networks using a modified susceptible-infected-susceptible (SIS) propagation mechanism. We also proposed a Non-Linear Dynamic System (NLDS)-based solution for the epidemic threshold, which determines the phase transition of the behavior of the system. This paper subsumes our previous work, and we briefly review the formulation and modeling of the problem for completeness. A key characteristic of our work is that it is applicable for any pair of topologies for the composite network. To validate our work, we use both synthetic composite networks with up to 50,000 of nodes, and a real composite network (specifically, mobile calls and text messages for an enterprise with 235 users for a given mobile carrier). In summary, our main contributions are:

1) We propose *EigenPredictor*, a method to predict which meme wins. The method relies on the insight from our analytical work mentioned above and uses multinomial logistic regression. A key strength of our method is that it can predict dominance by capturing the complex interplay of the two memes using just the first eigenvalues of appropriately defined system matrices for each meme. In extensive simulations with real and synthetic data, our method can predict the winner (dominating meme) with high accuracy, typically more than 95%.
2) We design and evaluate several suppression strategies based on the insights obtained from our analysis. We focus on two problems: (a) Unilateral Suppression, where we want to quench one meme while allowing the other to progress relatively unimpeded; and (b) Concurrent Suppression, in which we wish to eliminate both memes. Both suppression objectives are constrained by the number of allowed interventions, that is, the maximum number of nodes we can select to suppress the spread of the memes.
3) We evaluate the effect of cross-contamination between the layers of the composite network. In the cross-contamination scenarios, we allow the meme from one network to eventually "transform" into a meme on the other network. For example, think of a rumor propagating in Facebook being transformed into a rumor spread on twitter by an individual user. Cross-contamination adds a new dimension to the dynamic behavior of the competition, and this is arguably the first study that formulates and studies crossing-over.

The rest of the paper is organized as follows. We present the model and problem definitions in Section II and describe the proofs for the threshold in Section III. We present the EigenPredictor in Section IV. We show how to manage propagation and the effect of cross-contamination in section V. We discuss future directions in Section VI. In Section VII we review related work, and Section VIII concludes our paper.

II. Model and Problem Definitions

A. Our SI_1I_2S Propagation Model

We now present our meme propagation model for the competing memes; Table I explains our terminology. The propagation model is based on the popular "flu-like" SIS (Susceptible-Infected-Susceptible) model [8]. We name our model SI_1I_2S (Susceptible – Infected$_1$ – Infected$_2$ – Susceptible). Each node in the graph can be in one of three states: Susceptible (healthy), I_1 (infected by M_1), or I_2 (infected by M_2). The state transitions are shown in Fig. 1(b).

Meme persistence: δ. If a node is in state I_1 (or I_2), it recovers on its own with probability δ_1 (or δ_2). This parameter captures the persistence of the meme in an inverse way: a high δ means low persistence. Note that we assume that a node can only be infected by one meme, it cannot be infected by the other.

Meme strength: β. A healthy node gets infected by neighbors that got infected, and the meme strength is captured by β_1 and β_2. This potential infection is passed to a healthy neighbor in the absence of other interactions and we call this potential infection-in-isolation as an **attack**. In the following, we decide which infection succeeds (infects a susceptible node i) . Let C_1 be the number of attacks (each happening with probability β_1 independently) by node i's neighbors which are in state I_1 (infected by M_1); similarly, let C_2 be the number of neighbors infected by M_2. Then, we have three possible cases for a node in the Susceptible state:

- node i remains in the Susceptible state if $C_1 = 0$ and $C_2 = 0$.
- node i gets infected with M_1 with probability $\frac{C_1}{C_1+C_2}$.
- node i gets infected with M_2 with probability $\frac{C_2}{C_1+C_2}$.

B. Problem Definitions

Given the competing memes model described above, we now define the specific questions and problems that we address in this work. The input for all the problems is as follows:
Basic Input: (1) a composite network's adjacency matrices $\mathbf{A_1}$ and $\mathbf{A_2}$; and (2) the competing memes model's parameters: β_1, δ_1 for M_1 and β_2, δ_2 for M_2.

Given this basic input, we describe 4 problems, evenly divided into two broad categories: (1) *Dominance and Extinction* and (2) *Meme Control and Suppression*.

1) Dominance and Extinction: Our first two problems are motivated by the natural question of predicting the winner of the competition between the two memes.
Problem 1: Epidemic Threshold. Given: The basic input.
Find: A condition (threshold values for λ_1 and λ_2) under which the memes, either individually or collectively, die-out. This problem deals with the question of the epidemic threshold for our model, since under the threshold, neither meme survives.
Problem 2: Meme Dominance. Given: The basic input.
Find: Under the given set of parameters, which meme will dominate the system? Clearly, this question is more interesting in the case where one or both memes are above the epidemic threshold.
Defining dominance. Intuitively, we say that a meme dominates over the other meme if it manages to capture more

TABLE I
TERMINOLOGY

Symbol	Definition	Symbol	Definition
M_1, M_2	Meme #1, #2	$\mathbf{A_1}$, $\mathbf{A_2}$	Adjacency matrices
δ_1, δ_2	Meme persistence of M_1, M_2	β_1, β_2	Meme strength of M_1, M_2
S	Susceptible state	I_1, I_2	Infected state for M_1, M_2
$\mathbf{S_1}$, $\mathbf{S_2}$	System matrix for $\mathbf{A_1}$, $\mathbf{A_2}$, where $\mathbf{S} = (1-\delta)\mathbf{I} + \beta\mathbf{A}$	λ_1, λ_2	Largest eigenvalue of $\mathbf{S_1}$, $\mathbf{S_2}$ in absolute value.

nodes. The definition hides several subtleties, which have to do with the asymptotic behavior of the system, namely what happens as time goes to infinity [12]. However, as we are reliant on simulations, we are forced to adopt a more practical definition. First, we examine the behavior, after a sufficient warm-up period, when the system converges to some relatively stable state (with only small fluctuations of its infected nodes). We repeat our experiments 100 times and observe less than a 2% variance in the results. Assuming that we are beyond the warm-up period, we say that a meme prevails if it has infected at least θ% more network nodes than the other meme. For the results presented here we used $\theta = 10$, but we have also experimented with other θ values and observed similar qualitative results.

2) Meme Control and Suppression: The remaining problems involve our ability to control the outcome of the meme spread and competition.

Problem 3: Unilateral Suppression. Given: The basic input, and a subset of k nodes.

Find: The nodes to suppress, so that we maximize the chances that M_2 wins. Specifically, we want to find the best nodes to suppress meme M_1, so that we favor M_2. This problem is relevant in a market penetration of competing products, creating buzz in blogs, or a virus-versus-antivirus propagation problem in a computer or epidemiological setting.

Problem 4: Concurrent Suppression. Given: The basic input, and subset of k nodes.

Find: The best nodes to suppress that cause both memes to die-out. That is, we want to find the best set of nodes that reduce the spreading effectiveness of both memes. As an example of concurrent suppression, imagine the spread of two distinct, yet equally false rumors propagated through two distinct fringe pseudo-political groups that share memberships. Due to the potential damage such rumors may cause, both will be suppressed by preventing critical members from spreading the falsehoods.

III. THE EPIDEMIC THRESHOLD

In this section, We want to analytically determine the epidemic threshold (Problem 1). First, a discrete-time Non-Linear Dynamical System (NLDS), whose general form is $\mathbf{p}_{t+1} = g(\mathbf{p}_t)$, is used to approximate the infection process. The NLDS gives the evolution of the system with time, as we explain below. First, the probability that node i is infected by neighbor node j with meme M_1 at time t is $\beta_1 P_j^1(t-1)$. Then, the probability $\zeta_i^1(t)$ that node i does not receive the infection of M_1 from its neighbors (here, we assume the neighbors are independent) as:

$$\zeta_i^1(t) = \Pi_{j \in i's\ neighbors}(1 - \beta_1 P_j^1(t-1)) \quad (1)$$

Thus, we have the probability that node i receives the infection of M_1 at time t from its neighbors is:

$$1 - \zeta_i^1(t) = 1 - \Pi_{j \in i's\ neighbors}(1 - \beta_1 P_j^1(t-1)) \quad (2)$$

With the same reasoning, we can derive the probability of that node i receives the infection of M_2 from its neighbors at time t is:

$$1 - \zeta_i^2(t) = 1 - \Pi_{j \in i's\ neighbors}(1 - \beta_2 P_j^2(t-1)) \quad (3)$$

Now, we have the probability that node i is infected by M_1 from its neighbors at time t is the probability that node i receives the infection of M_1 and does not receive infection of M_2 from its neighbors at time t (assuming that the β and δ values are all extremely small, or, equivalently, the time between state transitions is extremely small.) Thus, we get:

$$T_i^1(t) = (1 - \zeta_i^1(t)) \cdot \zeta_i^2(t) \quad (4)$$

Using the same reasoning, we have the probability that the node is infected by M_2 at time t is:

$$T_i^2(t) = (1 - \zeta_i^2(t)) \cdot \zeta_i^1(t) \quad (5)$$

Hence the probability that node i is in state I_1 is:

$$P_i^1(t) = (1 - \delta_1) \cdot P_i^1(t-1) + T_i^1(t) \cdot S_i(t-1) \quad (6)$$

and the probability that it is in state I_2 is:

$$P_i^2(t) = (1 - \delta_2) \cdot P_i^2(t-1) + T_i^2(t) \cdot S_i(t-1) \quad (7)$$

and the probability that it is in state S (Susceptible) is:

$$S_i(t) = (1 - T_i^1(t) - T_i^2(t))S_i(t-1) + \delta_1 P_i^1(t-1) + \delta_2 P_i^2(t-1)$$

As mentioned before, for M_1 we define the vector $\vec{P}^1(t) = (P_1^1(t), P_2^1(t), ..., P_N^1(t))'$ where $P_i^1(t)$ is the probability that node i is infected by meme M_1 at time t. Similarly, for M_2, we have $\vec{P}^2(t) = (P_1^2(t), P_2^2(t), ..., P_N^2(t))'$. Let $\vec{V}(t) = (\vec{P}^1(t), \vec{P}^2(t))$ be the concatenation of two vectors. We use the NLDS formulation to describe the whole infection process evolution as $\vec{V}(t) = f(\vec{V}(t-1))$, with:

$$f_i(\vec{V}(t-1)) = \begin{cases} (1-\delta_1)P_i^1(t-1)+ \\ \quad T_i^1(t)S_i(t-1) & \text{if } i \le N \\ (1-\delta_2)P_i^2(t-1)+ \\ \quad T_i^2(t)S_i(t-1) & \text{if } i > N \end{cases} \quad (8)$$

Substituting $T_i^1(t)$, $T_i^1(t)$ and $S_i(t-1)$ into equation 8, we find that $f_i(\vec{V}(t-1))$ is equal to the following:

$$= \begin{cases} (1-\delta_1)P_i^1(t-1) + (1-\zeta_i^1(t))\zeta_i^2(t) \\ \quad (1-P_i^1(t-1) - P_i^2(t-1)) & \text{if } i \le N \\ (1-\delta_2)P_i^2(t-1) + (1-\zeta_i^2(t))\zeta_i^1(t) \\ \quad (1-P_i^1(t-1) - P_i^2(t-1)) & \text{if } i > N \end{cases}$$

We use the following theorem about the asymptotic stability of an NLDS at a fixed point:

Theorem 1 (Hirsch and Smale, 1974 [17]): The system given by $\mathbf{p}_{t+1} = g(\mathbf{p}_t)$ is asymptotically stable at an equilibrium point $\mathbf{p}^*$, if the eigenvalues of the Jacobian $J = \bigtriangledown g(\mathbf{p}^*)$ are less than 1 in absolute value, where

$$J_{k,l} = [\bigtriangledown g(\mathbf{p}^*)]_{k,l} = \frac{\partial p_{k,t+1}}{\partial p_{l,t}}|_{\mathbf{p}_t = \mathbf{p}^*}$$

The fixed point we are interested in for analyzing the threshold is the point where no node is infected (all nodes are healthy), i.e., $\vec{V}^* = \vec{0}$. By using this, we have the following theorem:

Theorem 2: The system is asymptotically stable at $\vec{V}^* = \vec{0}$ if the first eigenvalue of the system matrices for both memes as defined in Table I, are less than 1, i.e., $\lambda_1 < 1$ and $\lambda_2 < 1$, where λ_1 is the largest eigenvalue of matrix $S_1 = (1-\delta_1)I + \beta_1 A_1$ (and similarly for λ_2).

Proof: We are interested in the stability of the fixed point $\vec{V}^* = \vec{0}$. Let the Jacobian at this point be $\nabla(f)$ (a $2N$ x $2N$ matrix). Then

$$[\nabla(f)]_{ij} = \frac{\partial f_i(\vec{V}(t-1))}{\partial \vec{V}_j(t-1)}$$

Next, we can write it into a block matrix composed of the system matrices:

$$\nabla(f) = \left[\begin{array}{c|c} S_1 & S_3 \\ \hline S_4 & S_2 \end{array}\right]$$

In order to find the first eigenvalue of $\nabla(f)|_{\vec{V_f}}$, we define $\vec{X}$ as 2N elements vector:

$$\vec{X} = \left[\begin{array}{c} \vec{X}_1 \\ \hline \vec{X}_2 \end{array}\right]$$

where $\vec{X}_1$ and $\vec{X}_2$ have N elements respectively. We then have:

$$\bigtriangledown(f)|_{\vec{v_f}}\vec{X} = \left[\begin{array}{c|c} S_1 & S_3 \\ \hline S_4 & S_2 \end{array}\right] \cdot \left[\begin{array}{c} \vec{X}_1 \\ \vec{X}_2 \end{array}\right] = \lambda_{\bigtriangledown(f)}|_{\vec{v_f}} \left[\begin{array}{c} \vec{X}_1 \\ \vec{X}_2 \end{array}\right]$$

Doing the matrix multiplications, we get:

$$S_1\vec{X}_1 + S_3\vec{X}_2 = \lambda_{\bigtriangledown(f)}|_{\vec{v_f}}\vec{X}_1$$

and

$$S_4\vec{X}_1 + S_2\vec{X}_2 = \lambda_{\bigtriangledown(f)}|_{\vec{v_f}}\vec{X}_2$$

with $S_1 = (1-\delta_1)I + \beta_1 A_1$, $S_2 = (1-\delta_2)I + \beta_2 A_2$ and $S_3 = S_4 = 0$ (where I is the N x N identity matrix), as we show in Table I and as discussed below. Hence, the Jacobian $\nabla(f)$ is a block diagonal matrix and its eigenvalues are the same as the eigenvalues of S_1 and S_2. So the largest eigenvalue of $\nabla(f)$ can be either λ_1 or λ_2.

IV. EigenPredictor: Who Wins?

In this section, we determine which meme will prevail in the composite network, which we described as Problem 2. We showed in the previous section that when the system is below the threshold, both memes die-out. Hence, the use of the predictor is meaningful when the parameters are such that at least one of the memes is above threshold and Theorem 2 does not apply.

Due to the complexity of the problem, instead of an analytical solution, we present a predictive model, which we call *EigenPredictor*, which allows us to predict which meme will eventually prevail in the composite network. We also use simulations with synthetic and real data, which we describe below.

A. Simulation Set-up and Datasets

A discrete-time simulation of our system is used to simulate the stochastic behavior on different synthetic and real composite networks.

1) Small-scale Data sets ($N < 1,000$): **Real-world enterprise composite network (ENT).** This dataset, which was caputured over the course of six months, represents the phone call and SMS text message communications in one enterprise. Each node is an employee ($|N| = 235$), the edges in E_1 correspond to SMS messages exchanged between employees, and edges in E_2 correspond to phone calls made between employees. In detail, of all users, 31% communicate via calls alone, 28% via SMS alone, and 41% via both calls and SMS. In addition to the SMS and phone communication data, ENT also provides a basic social structure among the company's employees. This social information forms the basis of our `Social Hierarchy` method, which we will describe in Section §V. Each identified employee is grouped into one of 5 anonymous job roles. Job roles are ordered by their importance, thus we can determine who are the "bosses" and who are the regular employees [16].

Synthetic composite networks. Two synthetic graphs with 1,000 nodes are created: an Erdős-Rényi graph and a scale-free graph; we use the Barabási-Albert model [18]. Several different combinations of topologies have been experimented. Here, the reasons that we focus on these two are : (a)In order to show that our methods are not tailored to a particular family of graphs and (b) scale-free graphs are known to emerge in complex human and communication networks [18].

2) Large-scale Data Sets ($1,000 < N < 50,000$): In order to further stress-test the accuracy of our framework, we also did experiments on synthetic social networks with $1,000 < N < 50,000$ nodes that are generated by the `forestFire`,

`randomWalk`, and `nearestNeighbor` graph generation models [19], which are informed by real world measurements of social networks and provide graph structures that resemble such networks.

3) Simulation runs: We use a combination of Matlab and Python to conduct the experiments on real and synthetic composite networks. In each experiment, each meme infects a unique set of nodes Ini_1 and Ini_2. Each set of nodes has the same size, and is selected uniformly at random from N. The set of nodes is subject to the constraint $Ini_1 \cap Ini_2 = \emptyset$ (i.e., mutually exclusive). We run each simulation until it reaches a relatively stable state as we discussed in Section II. At which point, we determine the average number of nodes infected by M_1 and M_2 and report the outcome, which then gets averaged across 100 runs.

Accuracy. To measure the accuracy of our model, we compute, for each simulation, the percentage of runs where the outcome (as predicted by *EigenPredictor*) and the actual result (from the simulation) coincide.

B. The EigenPredictor

We now describe our EigenPredictor method for predicting which meme prevails. At stable state, we have three possible outcomes, which we represent as follows. If the outcome is '1', it means that M_1 will eventually prevail in the composite networks; if the outcome is '2', it means that M_2 will prevail; for the case where the difference is less than θ, which here is 10%, the outcome is '3'. Note that we experimented with other θ values (5%, 10%, and 15%) and the results were qualitatively similar. To summarize, given the parameters of our system, i.e., $A_1, \beta_1, \delta_1, A_2, \beta_2, \delta_2$, *EigenPredictor* produces a numeric value (1, 2, or 3), with the following semantics:

$$EigenPredictor((A_1,\beta_1,\delta_1),(A_2,\beta_2,\delta_2)) = \begin{cases} 1, & \text{if } M_1 \text{ prevails} \\ 2, & \text{if } M_2 \text{ prevails} \\ 3, & \text{if no clear winner} \end{cases} \quad (9)$$

Since λ_1 is the first eigenvalue (in absolute value) of system matrix S_1 for M_1, where $S_1 = (1-\delta_1)I + \beta_1 A_1$ (and similarly for λ_2), then λ_1 is a function of parameters A_1, β_1, δ_1 (and the same for λ_2). Therefore, equation 9 could be simplified and written as:

$$EigenPredictor(\lambda_1,\lambda_2) = \begin{cases} 1, & \text{if } M_1 \text{ prevails} \\ 2, & \text{if } M_2 \text{ prevails} \\ 3, & \text{if no clear winner} \end{cases} \quad (10)$$

We use our *EigenPredictor* in two scenarios: first, when λ_1 and/or λ_2 are below the threshold (Case 1); second, when both λ_1 and λ_2 are above the threshold (Case 2).

Case 1: At Least One Eigenvalue Below Threshold. From Section III, we know that if the system matrix's first eigenvalue of one meme is less than 1, the corresponding meme will die-out eventually. Therefore, in this scenario, we can predict which meme prevails eventually using the following three rules:

(i) if $\lambda_1 < 1$ and $\lambda_2 > 1$, then M_2 tends to prevail eventually in the composite networks;

(ii) if $\lambda_1 > 1$ and $\lambda_2 < 1$, then M_1 tends to prevail eventually in the composite networks;

(iii) if $\lambda_1 < 1$ and $\lambda_2 < 1$, then both memes will die out and none of them can be said to prevail.

Figures 2(a)-(e) demonstrate the proposed rules on both synthetic and real composite networks. The infection starts by infecting 30 nodes for each meme in Figure 2(a), Figure 2(b) and Figure 2(c), and 10 nodes for each meme in both Figure 2(d) and Figure 2(e). The outcomes of below- and above-threshold from these rules can be distinctly seen in these figures. These results show that, though simple, our proposed rules are very effective for predicting which meme tends to prevail eventually in the composite networks.

Case 2: Both Eigenvalues Above Threshold. This is the more interesting case in terms of competition: each meme in isolation would not die-out, so it is a "fight for dominance." We find again that the system eigenvalues play a critical role: the meme whose first eigenvalue is larger tends to prevail eventually in the composite networks.

Intuitively, the first system eigenvalues capture the likelihood of success for each meme. Recall that the system eigenvalue considers both the topology and the meme strength. Extensive experimental results, shown in Figure 2(f), argue in favor of this observation: we plot the outcome of the competition for different pairs of the eigenvalues (λ_2 vs. λ_1); we see that above the diagonal M_2 prevails, and below the diagonal M_1 prevails, in other words, the meme with the largest system eigenvalue wins.

Our regression model. We verify the significance of λ_1 and λ_2 as determining factors using a rigorous regression model. In particular, we use a multinomial logistic regression [20] to predict the outcome in the case where both eigenvalues are above the threshold:

$$log\frac{Pr(Y=i|X=(\lambda_1,\lambda_2))}{Pr(Y=K|X=(\lambda_1,\lambda_2))} = \alpha_{i0} + \alpha_{i1}\cdot\lambda_1 + \alpha_{i2}\cdot\lambda_2 \quad (11)$$

Here Y represents the outcome of this model (dependent variable), and X represents the input, i.e., the independent variables λ_1 and λ_2; $\vec{\alpha_i}$=(α_{i0}, α_{i1} , α_{i2}) is the coefficient vector of independent variables, where i=1,2. Therefore, we have:

$$Pr(Y=i|X=(\lambda_1,\lambda_2)) = \frac{exp(\alpha_{i0}+\alpha_{i1}\cdot\lambda_1+\alpha_{i2}\cdot\lambda_2)}{1+\sum_{i=1}^{K-1} exp(\alpha_{i0}+\alpha_{i1}\cdot\lambda_1+\alpha_{i2}\cdot\lambda_2)} \quad (12)$$

$$Pr(Y=K|X=(\lambda_1,\lambda_2)) = \frac{1}{1+\sum_{i=1}^{K-1} exp(\alpha_{i0}+\alpha_{i1}\cdot\lambda_1+\alpha_{i2}\cdot\lambda_2)} \quad (13)$$

where i=1,2 and K=3. The coefficients vectors α_i computed by our regression model are presented in Table III; note that corresponding coefficients' values (e.g., α_{10} and α_{20}) are not, as one might expect at first, equal in magnitude and of opposite signs, since the output range is 1–3.

To construct the ground truth, we ran system simulations (input: λ_1 and λ_2; output: which meme prevails) on both synthetic and real composite networks. The number of data points was 5,339 for synthetic composite networks and 6,844 data points for real composite networks (see Section IV-A for

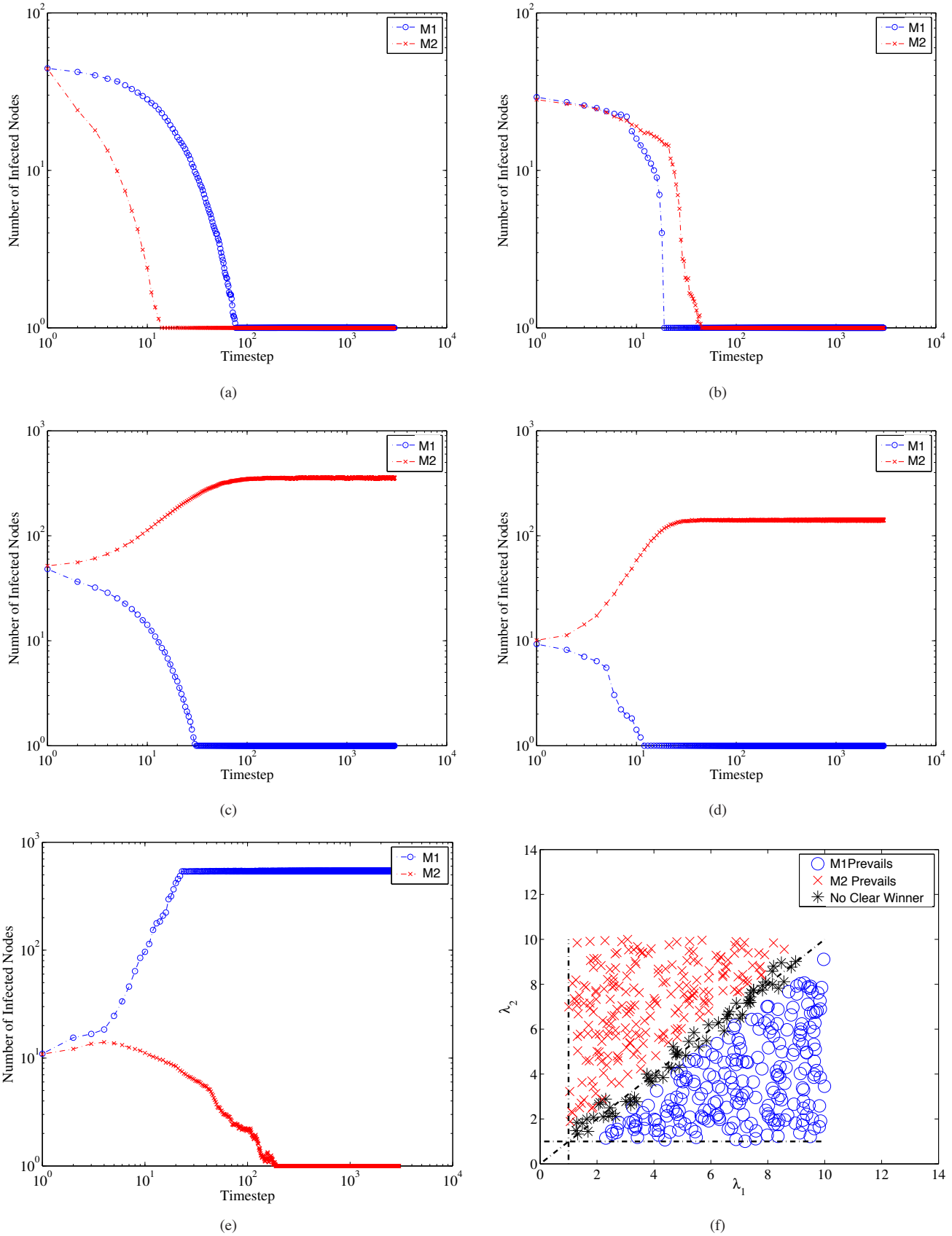


Fig. 2. Simulation Results: Infection plot over time (log-log) in Figure(a)-(e). 2(a): Synthetic Composite Networks: $\lambda_1 = 0.97, \lambda_2 = 0.96$; 2(b): Real Composite Networks: $\lambda_1 = 0.9, \lambda_2 = 0.94$; 2(c): Synthetic Composite Networks: $\lambda_1 = 0.91, \lambda_2 = 1.63$; 2(d): Real Composite Networks: $\lambda_1 = 0.99, \lambda_2 = 1.4$; 2(e): $\lambda_1 = 4.5, \lambda_2 = 1.7$; 2(f): The outcomes for different combinations of system eigenvalues: $1 < \lambda_1 < 10$ and $1 < \lambda_2 < 10$; black dotted lines represent three lines λ_1=1, λ_2=1, and λ_1=λ_2. When the eigenvalues are roughly equal there is no clear winner.

TABLE II
PREDICTION ACCURACY.

Training data set size	Accuracy	
	Synthetic networks	Real networks
5%	96.65%	95.05%
10%	98.42%	98.28%

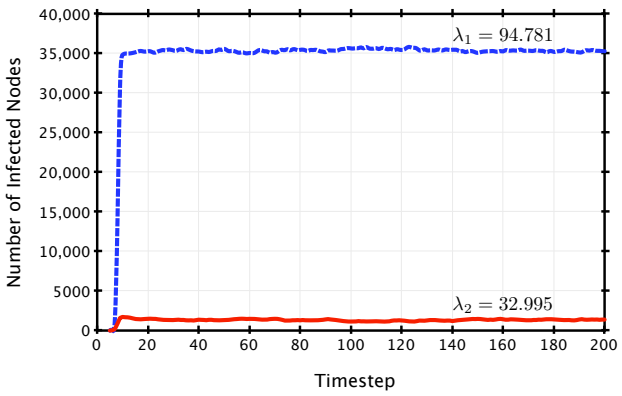
a description of the datasets). Next, for each data point, we compare the outcome from the simulation with the outcome from the *EigenPredictor* (regression), and compute the accuracy of the predictor.

As shown in Table III, using only 5% of the dataset as training data (and 95% of the dataset for validation), our model achieves a prediction accuracy of 96.65%; when using 10% for training, the accuracy is 98.42%, with 95% confidence interval.

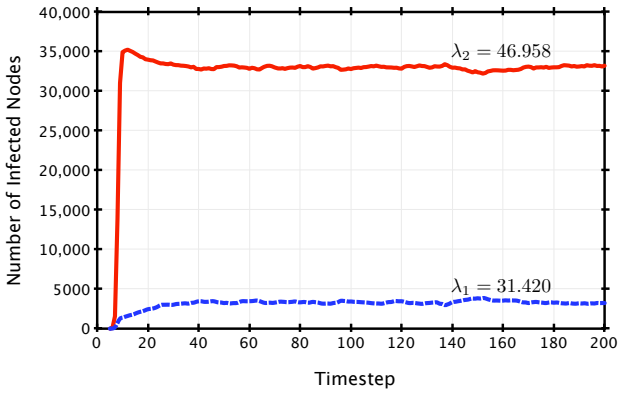
Stress-testing EigenPredictor: topology independence. In order to test the robustness of our model, we use a different training topology from the testing topology. Specifically, we use 5% and 10% of the data set from *synthetic* composite networks as training data to predict the data set from *real* composite networks. We achieve an accuracy of 95.05% and 98.28%, respectively, with a 95% confidence interval, as shown in Table II. These results indicate that *EigenPredictor* is very effective in practice.

We now present details on the statistical results and significance of our model: Table III shows the results of the logistic regression ($\vec{\alpha}_1$, $\vec{\alpha}_2$, and $|t|$-values). To understand how well the independent variables λ_1 and λ_2 contribute to the model (i.e., explain the dependent variable Y), note the $|t|$-values; since the $|t|$-values are much larger than 2.08 (which translates to p-values being much lower than < 0.01), this indicates that our model is statistically significant at the 1% level. These results indicate that all features contribute to the prediction model, and the model has high predictive power.

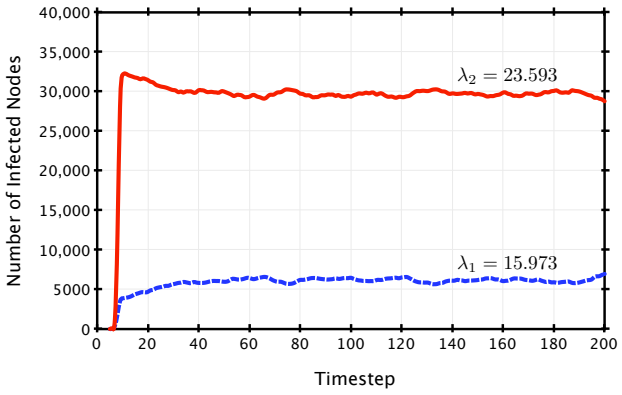
Discussion. The significance of the multinomial logistic regression is two-fold: (a) we verify quantitatively that the λ_1 and λ_2 are statistically sufficient to determine the winning meme and (b) we develop a practical method to predict meme dominance under complex dynamics.



(a) $\lambda_1 > \lambda_2$



(b) $\lambda_1 < \lambda_2$



(c) $\lambda_1 < \lambda_2$

Fig. 3. Large Scale Experimental Results. This figure shows experimental results of large scale epidemic simulations using a ForestFire and Nearest Neighbor synthetic graph models. Results above are for $N = 40,000$ nodes, but are consistent for results of $10,000$ to $50,000$ node experiments. Unlike smaller-scale experiments, these results show that the weaker meme may retain some endemic population, yet the meme with the larger eigenvalue clearly dominates the simulation.

V. MEME SUPPRESSION, CONTROL AND CROSS-CONTAMINATION

In this section, we first design and evaluate suppression methods based on two distinct strategies (described as problems 3 and 4 in §II-B2):

1) *Unilateral Suppression*. The goal of this strategy is to suppress one meme, while leaving the other unscathed, thus free to spread unimpeded. Using the five techniques described below, we intend to suppress M_1 by removing ability to spread the memes from a subset of carefully selected nodes. The intention of unilateral suppression is to reduce λ_1 to below λ_2 (i.e., $\lambda_1 < \lambda_2$), thus affecting the outcome.
2) *Concurrent Suppression*. The goal of this strategy is to suppress the spread of both memes by removing a set of nodes from both graphs in the composite network. Ultimately, we want to reduce λ_1 and λ_2 to below 1 (i.e., $\lambda_1, \lambda_2 < 1$), thus stopping the spread of both memes.

Given these two main strategies, we propose 5 methods, described in Table IV, partially motivated by the methods used in single virus/disease propagation on a single network [21], [22]. We evaluate each method's effect on

TABLE III
REGRESSION RESULTS: COEFFICIENT VECTORS AND THEIR CORRESPONDING $|t|$-VALUES.

	α_{i0}	α_{i1}	α_{i2}
$\vec{\alpha}_1$	-3.43	1.44	-1.32
$\|t\|$-values for $\vec{\alpha}_1$	9.27	12.18	8.66
$\vec{\alpha}_2$	7.05	1.37	2.39
$\|t\|$-values for $\vec{\alpha}_2$	6.64	9.64	12.48

the system matrix eigenvalues for each subgraph in the composite network (λ_1,λ_2). The proposed methods are: (a) `Random`, (b) `Acquaintance`, (c) `Max Degree`, (d) `Social Hierarchy` and (e) `Greedy`.

A. Unilateral Suppression

As mentioned above, the objective of unilateral suppression is to reduce λ_1 to less than λ_2, thus reversing the prediction of our EigenPredictor. That is, we seek to answer: *What set of nodes should we suppress in order to reduce the spread of one meme, ultimately resulting in the dominance of the other, unsuppressed, meme?* We present the results of using Unilateral Suppression on the enterprise data set in Figure 4. Note that M_1 will eventually prevail in the composite network prior to applying any unilateral suppression strategies. Then, observe that the value of λ_1 decreases as nodes are removed from the system. At $k = 10$, λ_1 is reduced to below λ_2 (thus reversing the prediction of the *EigenPredictor*).

As expected, the two methods that rely on randomness (i.e., `Random` and `Acquaintance`) have the worst performance compared to the other methods. In contrast, `Greedy` performs better than the others, yet is the most expensive computationally. `Max Degree` performs surprisingly well, within 1% of `Greedy` at much lower computational cost.

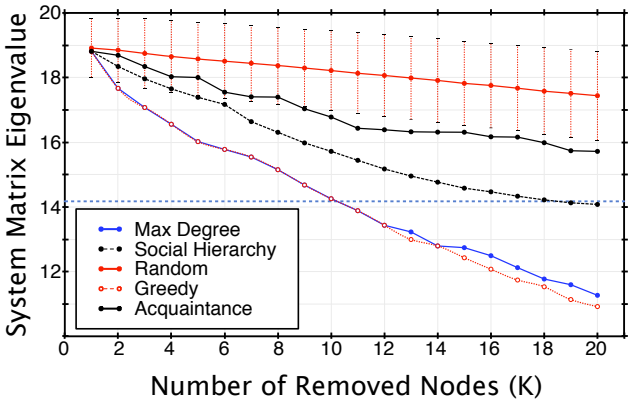
Interestingly, when we remove nodes based on their social status (e.g., remove "bosses" before "managers," and so on), the method `Social Hierarchy` performs better than the random methods, yet not as well as the topologically-informed models, and eventually (at $k = 20$) crosses the value of λ_2. Though not as effective, in situations where we lack topological information, we could potentially rely on easily observable social hierarchy information to inform our suppression process.

B. Concurrent Suppression

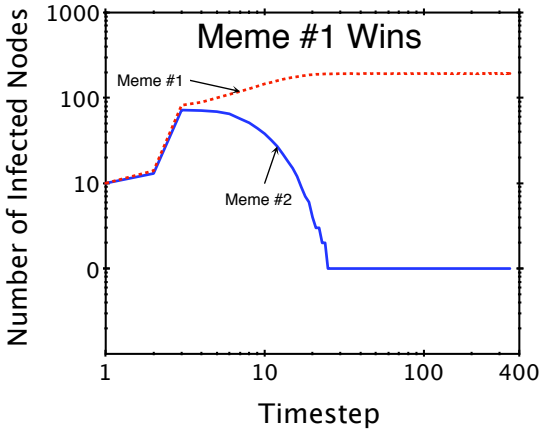
Under the concurrent suppression scheme, the goal is to reduce the effective spreading power of two memes spreading through different modes of communication (i.e., edge sets E_1, E_2 in a composite network). Simply put, we ask: *What set of shared nodes should we inoculate in order to reduce the spread of both memes the most?*

We present the results of our suppression methods in Figure 5. As before, we observe that `Max Degree` and `Greedy` reduce both λ_1 and λ_2 to below the epidemic threshold (indicate by the horizontal line at 1) at approximately $k = 19$.

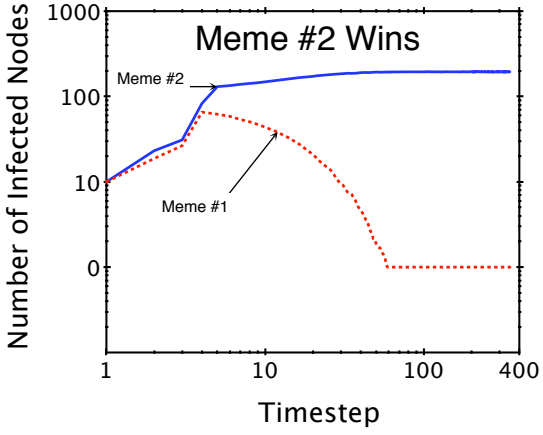
`Social Hierarchy` provides mixed results. As indicated by the lower plotted line (representing the SMS portion of the enterprise network), `Social Hierarchy` is nearly as effective as `Max Degree` and `Greedy`. Yet, on the upper line (call graph of the enterprise network), the `Social Hierarchy` method is not as effective and does not reduce the spreading power to below the epidemic threshold.



(a)



(b) Before Suppression



(c) After Suppression

Fig. 4. Example of Unilateral Suppression on the enterprise data set. The methods `Greedy`, `Max Degree` and `Social Hierarchy` drop the system matrix eigenvalue λ_1 below λ_2 (thus reversing the prediction of the EigenPredictor); (b) shows the original competition results without removing nodes; note that M_1 wins, while M_2 dies out; (c) shows the competition results after removing $k = 20$ nodes using the `Max Degree` method; the result is reversed, with M_2 winning and M_1 dying out.

TABLE IV
SUPPRESSION METHODS

Method Name	Unilateral	Concurrent	Intuition
`Random`	$rand(node_{G_1})$	$rand(node_{G_1\|G_2})$	Randomly select a node and remove it from G_1 ($G_1\|G_2$)
`Acquaintance`	$rand(neighbor)$ of $rand(node)$	$rand(neighbor)$ of $rand(node)$ of $rand(G_1\|G_2)$	Acquaintance immunization, remove a random neighbor of a randomly selected node in G_1 ($G_1\|G_2$).
`Max Degree`	$max(deg(G_1))$	$max(deg(G_1\|G_2))$	Remove node with the maximum degree in G_1 ($G_1\|G_2$).
`Social Hierarchy`	$max(rank(node))$	$max(rank(node))$	Remove node with the highest rank.
`Greedy`	$max(\lambda_1)$	$max(\lambda_1\|\lambda_2)$	Remove the node that causes the largest eigenvalue drop in either λ_1 or λ_2.

C. Summary

In summary, we have designed and evaluated several techniques for unilateral and concurrent suppressions, which are based on randomness, topological information and social hierarchy. The results from both suppressions show that the topological properties-based method (i.e., `Max Degree`) is very effective in controlling meme propagation compared to other methods. Put another way, removing the highest-connected node is a very effective suppression strategy. In situations where we lack topological information, we could potentially rely on the explicit information of social hierarchy to design our suppression scheme (e.g., `Social Hierarchy`), though not as effective as the topological properties-based method.

D. Cross-Contamination Experiments

Until this point, we have considered memes spreading on a composite network to be mutually exclusive, i.e., a meme will spread using only the edges associated with their own network. In this section, we evaluate the effect of cross-contamination across various synthetic graphs. We change the model slightly: we allow the meme from one network to eventually "transform" into a meme that can propagate on the edges of the other network. For example, consider a rumor propagating in Facebook being transformed into a rumor spread on twitter by an individual user. The user creates a new meme and releases it on twitter, which spreads the same information, but now this meme is spreading across twitter edges. Our simulation model emulates this exact scenario, by carefully following the propagation of these "cross-over" meme. To account for the new ability of a meme to jump composite network layers, each meme is assigned a cross-contamination parameter, denoted $0 < X_{A\to B}, X_{B\to A} \leq 1.0$. Specifically, $X_{A\to B}$ describes the ability of a meme propagating on composite network layer A to cross to layer B. $X_{A\to B}$ is similarly defined.

In the shown scenarios, if simulated in isolation, each of the memes would propagate and capture the graph. While in competition, in each of these simulations, ultimately one meme dominates. We have established that the meme propagating with the largest eigenvalue will eventually dominate the graph in the absence of cross-over. However, the crossing over allows the meme with higher such likelihood to spill over and propagate on the other topology, thus giving an advantage to the meme with higher such likelihood.

In Figure 6(a), we demonstrate the results of cross contamination on a composite network of 500 nodes. In this example, $\lambda_{S,1} = 26.43$, and $\lambda_{S,2} = 13.24$ with cross-contamination probabilities of $X_{A\to B} = .15$ and $X_{B\to A} = 0.05$. Observing that $\lambda_{S,1} > \lambda_{S_2}$, we see that meme #1 ultimately captures the greatest number of nodes. Interestingly, meme #2 does not completely die out, possible due to the ability of meme #2 to cross over to both layers of the composite network.

In Figure 6(b), we show the average steady-state percentage of infected nodes versus the propagations strength of meme #2 (β_2). All other parameters are constant, in particular $\beta_1 = 0.25$, $X_{A\to B} = 0.10$, and $X_{B\to A} = 0.10$. We observe that meme #2 will eventually dominate the simulation as we increase β_2. Due to space limitations, we cannot present a more extensive study of this case, despite having more results, while it would be interesting to study this problem analytically.

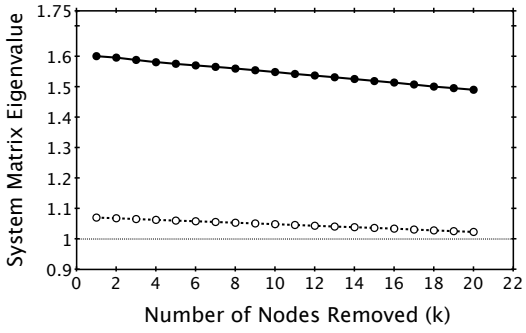
VI. DISCUSSION AND FUTURE WORK

In this section, we discuss the limitations of our work and possible future directions.

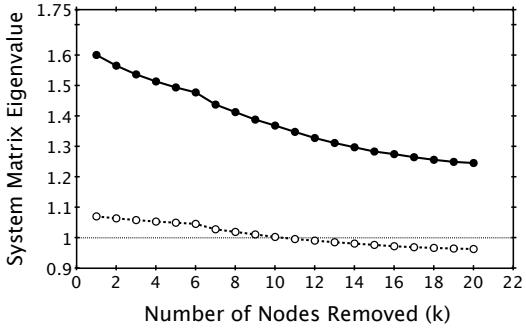
Choice of epidemic model. The flu-like SIS (Susceptible-Infected-Susceptible) epidemiological model is simple, yet illustrative, and has been extensively studied in past literature in a single-virus setting (cf. [9], [1], [23]). Therefore, we chose to extend SIS in order to gain fundamental insights into the dynamics of competing memes. We leave the investigation of other epidemic models as future work.

Using real composite networks. Finding real data for any networked system or communication is non-trivial due to privacy concerns, infrastructure limitations, and measurement biases. Finding real data sets of *composite networks* is even more challenging. Obtaining the enterprise dataset used in this paper was instrumental in modeling and understanding how real composite network operate, but the dataset comes with dissemination restrictions. We believe that the research community could greatly benefit from the creation of an open repository of real composite networks.

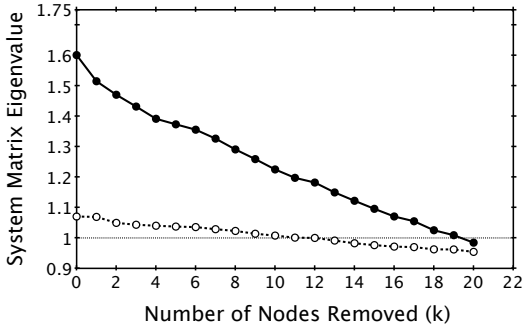
Deeper exploration. Our paper is the first attempt to study, predict and manage competing epidemic propagations on composite networks. We leave further exploration, like finding the extent of foot-prints, proving performance bounds for inoculation policies and incorporating more elaborate interactions between the two networks as future work.



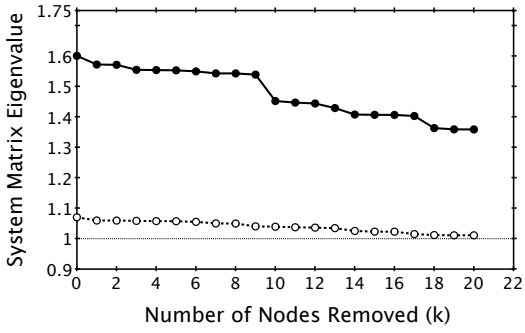
(a) Random



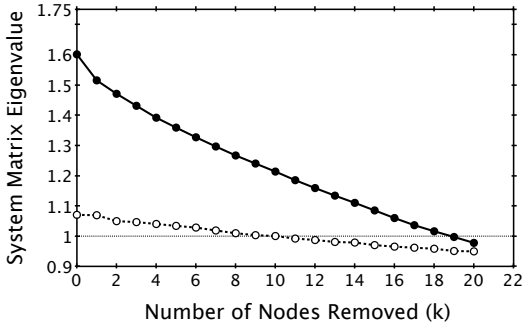
(b) Social Hierarchy



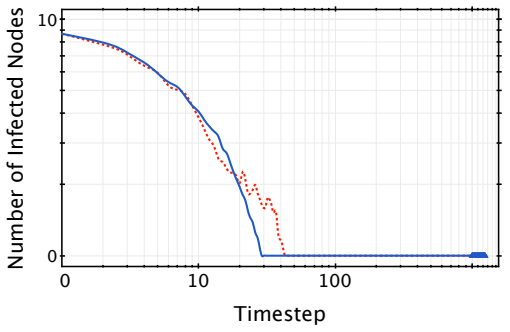
(c) Max Degree



(d) Acquaintance

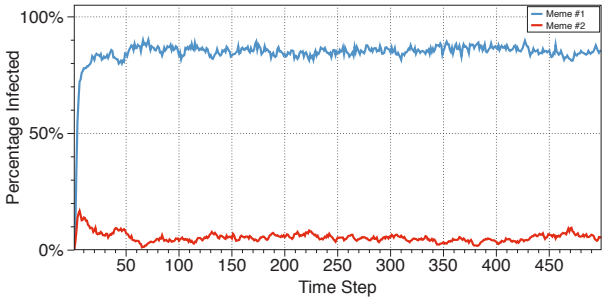


(e) Greedy

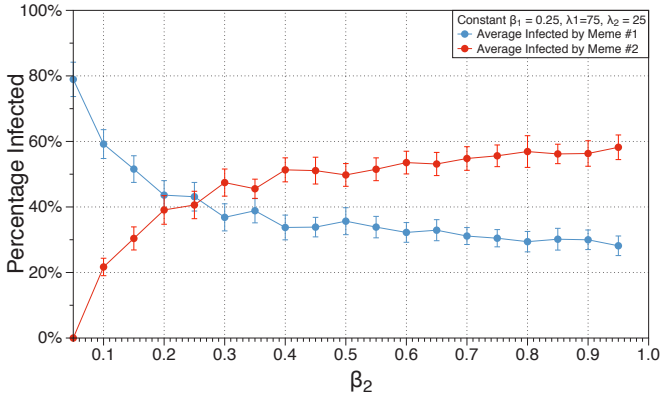


(f) Outcome of Concurrent Suppression

Fig. 5. Example of Concurrent Suppression on the enterprise data set, using each method. The epidemic threshold is marked in each plot at 1. Again, both the Greedy and Max Degree methods drop λ_1 and λ_2 below the epidemic threshold. Subplot 5(f) shows suppression results after removing $k = 20$ nodes selected using the Max Degree method—both memes die out.



(a) Example of Cross-Contamination



(b) Cross-Contamination vs. β_2

Fig. 6. Cross contamination simulation results.

VII. Related Work

We now proceed to reviewing related work in the context of single-meme and multiple-meme propagation from epidemiology, communication networks, game theory, and data mining literature.

Single-meme propagation. Many works focus on single meme propagation on one single topology. Compartmental

models like SIS, SIR, etc., have been well-studied in many epidemiological texts [9], [1], [24]. The evolution of blogs and the maximization of influence propagation are studied in [3], [25]. Information cascades models are proposed to study the meme propagation in word-of-mouth communications [26], [2]. Richardson et al. proposed a meme propagation model to achieve optimal viral marketing plans [27]. Numerous studies exist on virus propagation on the Internet based on the basic epidemic models of infection [4], [5]. Virus propagation under special cases have been studied, e.g., in IPv6 Internet [28]. A fundamental question in epidemiology is the presence of a threshold, under which an epidemic is guaranteed not to happen. Pastor-Satorras et al. [10] proposed an epidemic threshold condition for random power law networks, which uses the "mean-field" approach. Ganesh et al. [13] and Yang et al. [11] provided epidemic threshold for the single-virus on single topology. Prakash et al. [12] gave the epidemic threshold condition for almost all single-virus epidemic models on a single static network. Cohen et al. [21] studied the well-known acquaintance immunization method and showed that it is much better than random methods. Tong et al. [22] provide a simple greedy $(1 - 1/e)$ approximation algorithm for immunizing nodes under the SIS model.

Multiple memes and interdependent networks. Newman [14] studied multiple viruses on a single, special random graph and provided the epidemic threshold for the case when the second virus propagates over the residual network after the propagation of the first virus has completed. This scenario is close to the dynamics of propagation of a single virus—one virus passed over the network, the second virus starts to pass over the residual network. Models for multiple cascades have been studied as extensions of the independent cascade model, where once a node is infected with a cascade, it never change its state [29]. Multiple viruses propagation on simple fair-play single network was investigated [30]. The effects of cascades in inter-dependent networks (e.g., Internet router and power electricity networks) were investigated by Buldyrev et al. [31]. However, all of these works are completely different from our problem as we consider the more realistic and challenging scenario of competing memes propagating simultaneously on composite networks. In this paper, we have significantly extended our preliminary work [16] by: (a) proposing and evaluating an effective prediction scheme, Eigenpredictor, (b) considering immunization strategies with various suppression techniques to affect the outcome of the propagation, and (c) evaluating the effect of cross-contamination across various graphs.

Game theory. Meier et al. [32] studied inoculation games in social networks, where each node selfishly decides whether or not to protect itself. The game between a virus and an alert over a network was investigated by Aspnes et al. [33]. Kostka et al. [34] studied competing campaigns as a game-theoretical problem and showed that being the first player was not always advantageous. However, these works using game theory are different from our problem where we assume that all nodes are passive and follow the same propagation model.

VIII. Conclusion

In this paper, we have designed an effective methodology, *EigenPredictor*, to predict which meme will eventually prevail. Our theoretical and experimental results show that *EigenPredictor* achieves very high accuracy (above 95%) on a wide variety of real and synthetic datasets. Given the outcomes predicted by our *EigenPredictor*, we have designed and evaluated various suppression schemes to alter the results of the competing memes on composite networks. Extensive experimental results have revealed the comparative effectiveness of suppression schemes. Finally, we formulate and provide an initial study of the effect of cross-contamination across the composite graphs, where a meme cross-over and starts propagating on the other network's topology.

Acknowledgements

We would like to thank the anonymous reviewers for their feedback. This material is based upon work supported by the National Science Foundation under Grants No.IIS1017415, CNS-1064646 and NSF CISE NECO-0832069, and by DARPA SMISC W911NF-12-C-0028 and ARL W911NF-09-2-0053. This work is also partially supported by funds from the VT College of Engineering.

References

[1] A. G. McKendrick, "Applications of mathematics to medical problems," *Edin. Math. Society*, 1926.
[2] 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*, 2001.
[3] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, "Information diffusion through blogspace," in *WWW*, 2004.
[4] S. Staniford, V. Paxson, and N. Weaver, "How to 0wn the Internet in Your Spare Time," in *USENIX Security Symposium*, 2002.
[5] R. W. Thommes and M. J. Coates, "Epidemiological Modelling of Peer-to-Peer Viruses and Pollution," in *IEEE INFOCOM*, 2006.
[6] E. M. Rogers, *Diffusion of Innovations, 5th Edition*. Free Press, August 2003.
[7] N. Ungerleider, "Massive Egyptian Protests Powerd by YouTube, Twitter, Facebook, Twitpic," in *Fast Company*, 2011.
[8] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, 2000.
[9] R. M. Anderson and R. M. May, *Infectious diseases of humans: Dynamics and control*. Oxford Press, 2002.
[10] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics in finite size scale-free networks," *Phys. Rev. E*, vol. 65, 2002.
[11] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: an eigenvalue viewpoint," in *IEEE SRDS*, 2003.
[12] B. A. Prakash, D. Chakrabarti, M. Faloutsos, N. Valler, and C. Faloutsos, "Threshold conditions for arbitrary cascade models on arbitrary networks," *IEEE ICDM*, 2011.
[13] A. Ganesh, L. Massoulie, and D. Towsley, "The Effect of Network Topology in Spread of Epidemics," in *IEEE INFOCOM*, 2005.
[14] M. E. J. Newman, "Threshold effects for two pathogens spreading on a network," *Physical Review Letters*, 2005.
[15] F. Li, Y. Yang, and J. Wu, "CPMC: An Efficient Proximity Malware Coping Scheme in Smartphone-based Mobile Networks," in *IEEE INFOCOM*, 2010.
[16] X. Wei, N. Valler, B. A. Prakash, I. Neamtiu, M. Faloutsos, and C. Faloutsos, "Competing Meme Propagation on Networks: A Case Study of Composite Networks," in *ACM CCR*, 2012.
[17] M. W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, 1974.
[18] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, 1999.
[19] A. Sala, L. Cao, C. Wilson, R. Zablit, H. Zheng, and B. Y. Zhao, "Measurement-calibrated graph models for social network experiments," in *WWW*, 2010.

[20] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*. John Wiley, 1989.
[21] R. Cohen, S. Havlin, and D. ben Avraham, "Efficient immunization strategies for computer networks and populations," *Physical Review Letters*, 2003.
[22] H. Tong, B. A. Prakash, C. E. Tsourakakis, T. Eliassi-Rad, C. Faloutsos, and D. H. Chau, "On the vulnerability of large graphs," in *IEEE ICDM*, 2010.
[23] J. O. Kephart and S. R. White, "Directed-Graph Epidemiological Models of Computer Viruses," *IEEE S&P*, 1991.
[24] ——, "Measuring and modeling computer virus prevalence," in *IEEE S&P*, 1993.
[25] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *ACM SIGKDD*, 2003.
[26] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change in informational cascades," *J. Political Economy*, 1992.
[27] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *ACM SIGKDD*, 2002.
[28] Z. Chen and C. Ji, "Measuring Network-Aware Worm Spreading Ability," in *IEEE INFOCOM*, 2007.
[29] C. Budak, D. Agrawal, and A. E. Abbadi, "Limiting the Spread of Misinformation in Social Networks," in *WWW*, 2011.
[30] B. A. Prakash, A. Beutel and C. Faloutsos., "Winner-takes-all: Competing Viruses on fair-play networks," in *WWW*, 2012.
[31] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, 2010.
[32] D. Meier, Y. A. Oswald, S. Schmid, and R. Wattenhofer, "On the windfall of friendship: inoculation strategies on social networks," in *ACM EC*, 2008.
[33] J. Aspnes, N. Rustagi, and J. Saia, "Worm versus alert: Who wins in a battle for control of a large-scale network?" in *OPODIS*, 2007.
[34] J. Kostka, Y. Oswald, and R. Wattenhofer, "Word of mouth: Rumor dissemination in social networks," in *SIROCCO*, 2008.

Xuetao Wei is a Ph.D candidate in the department of computer science and engineering at University of California, Riverside. His research interests span the areas of systems, security and networks with an emphasis on smartphones, network science, and green networks and systems.

Nicholas Valler has received his Ph.D. from the University of California, Riverside in the Summer of 2012. His research interests include network science, network operations, and data mining applications. He currently works for CrowdCompass, Inc., in Portland, Oregon as a DevOps Engineer.

B. Aditya Prakash is an Assistant Professor in the Computer Science Department at Virginia Tech. He graduated with a Ph.D. from the Computer Science Department at Carnegie Mellon University in 2012, and got his B.Tech (in CS) from the Indian Institute of Technology (IIT) - Bombay in 2007. He has published 22 refereed papers in major venues and holds two U.S. patents and has given two tutorials (VLDB 2012 and ECML/PKDD 2012). His work has received one best paper award and two best-of-conference selections (CIKM 2012, ICDM 2012, ICDM 2011). His interests include Data Mining, Applied Machine Learning and Databases, with emphasis on large real-world networks and time-series.

Iulian Neamtiu is an assistant professor in the Department of Computer Science and Engineering at the University of California, Riverside. He received his Ph.D. from the University of Maryland, College Park in 2008. His research interests are in programming languages, software engineering, and the smartphone side of systems and security. He is a recipient of the NSF CAREER award.

Michalis Faloutsos is a faculty member of the Computer Science Dpt in the University of New Mexico since 2012. He got his bachelor's degree at the National Technical University of Athens and his M.Sc and Ph.D. at the University of Toronto. His interests include, Internet protocols and measurements, peer-to-peer networks, network security, BGP routing, and ad-hoc networks. With his two brothers, he co-authored the paper on power-laws of the Internet topology, which received the ACM SIGCOMM Test of Time award. His work has been supported by many NSF and military grants, for a cumulative total of more than $6 million. Several recent works have been widely cited in popular printed and electronic press such as slashdot, ACM Electronic News, USA Today, and Wired. Most recently he has focused on the classification of traffic and web-security, and co-founded a cyber-security company in 2008, offering services as www.stopthehacker.com, which received two SBIR grants from the National Science Foundation, and institutional funding in Dec 2011.

Christos Faloutsos is a Professor at Carnegie Mellon University. He has received the Presidential Young Investigator Award by the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, the SIGKDD Innovations Award (2010), nineteen "best paper" awards (including two "test of time" awards), and four teaching awards. He is an ACM Fellow, he has served as a member of the executive committee of SIGKDD; he has published over 200 refereed articles, 11 book chapters and one monograph. He holds six patents and he has given over 30 tutorials and over 10 invited distinguished lectures. His research interests include data mining for graphs and streams, fractals, database performance, and indexing for multimedia and bio-informatics data.