On the Power of Adversarial Infections in Networks

Michael Brautbar[1] *, Moez Draief[2], and Sanjeev Khanna[1]

[1] Computer and Information Science, University of Pennsylvania
brautbar@cis.upenn.edu , sanjeev@cis.upenn.edu
[2] Electrical and Electronic Engineering, Imperial College London
m.draief@imperial.ac.uk

Abstract. Over the last decade we have witnessed the rapid proliferation of online networks and Internet activity. Such activity is considered as a blessing but it brings with it a large increase in risk of computer malware — malignant software that actively spreads from one computer to another. To date, the majority of existing models of malware spread use stochastic behavior, when the set of neighbors infected from the current set of infected nodes is chosen obliviously. In this work, we initiate the study of *adversarial* infection strategies which can decide intelligently which neighbors of infected nodes to infect next in order to maximize their spread, while maintaining a similar "signature" as the oblivious stochastic infection strategy as not to be discovered. We first establish that computing an optimal and near-optimal adversarial strategies is computational hard. We then identify necessary and sufficient conditions in terms of network structure and edge infection probabilities such that the adversarial process can infect polynomially more nodes than the stochastic process while maintaining a similar "signature" as the oblivious stochastic infection strategy. Among our results is a surprising connection between an additional structural quantity of interest in a network, the network *toughness*, and adversarial infections. Based on the network toughness, we characterize networks where existence of adversarial strategies that are pandemic (infect all nodes) is guaranteed, as well as efficiently computable.

1 Introduction

Over the last decade we have witnessed the rapid proliferation of online networks and Internet activity. While such a proliferation is considered by many as a blessing, it brings with it an increase in risk of computer malware — malignant software that actively spreads from one computer to another. Indeed, a long thread of research has been devoted to understanding the spread of malicious malware such as computer viruses, computer worms and other malignant forms of computer infection cf. [3, 16, 17, 19, 21]. However, such work has, so far,

[*] Now in The Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, brautbar@mit.edu.

only considered oblivious malware propagation, where the spreading malware does not behave strategically in its choice of which nodes to spread to. What now is a standard way of containing the spread of malware is the control of the amount of information that spreads from one computer to others [3, 19]. This is also known as throttling. Under rate control, a malware that does not want to get detected spreads obliviously to a small set of neighbors while abiding by the rate constraint. Inspired by this fact, in this work we initiate the study of *intelligent* malware propagation, where the spreading malware can strategically decide which neighboring nodes to infect under rate constraints, in order to maximize the total number of infections over time. Each edge (u, v) is equipped with an edge weight $p(u, v)$ representing the amount of a typical and normal communication between the two nodes u and v. Typical examples of such networks include email networks, instant messaging networks, and online social networks, among others. Under the rate constraint, a malware spreading from an infected node u can infect at most a number of nodes that is not more than the typical "signature" of communication, namely,

$$\left\lceil \sum_{\{v \text{ neighbor of } u\}} p(u,v) \right\rceil .$$

More generally, we will demand that for any subset X, the malware must not infect more nodes than the amount of traffic involving X permits it to, namely,

$$\left\lceil \sum_{u \in X} \sum_{\{v \text{ neighbor of } u\}} p(u,v) \right\rceil .$$

In particular, such malware can use the structure of the network in order to choose which of the neighboring computers to infect from a newly infected node. We initiate a detailed comparison of the behavior of an adversarial infection, that can use the network structure, to that of an oblivious stochastic one, that spreads according to the standard Independent Cascades diffusion model. In order to defend against future malware it is of primary importance to first analyse the adversarial infections strategy and to contrast its behavior with that of the oblivious stochastic strategy.

We would like to further emphasize the need to understand the behavior of adversarial infections with the following example, comparing the behavior of a well-planned adversarial infection to that of a simple heuristic. Consider a path connected at one of its ends to the root of a two level binary tree. Set all edge weights to 1/2. A greedy strategy may consist, for each newly infected node u, to infect its r_u neighbors of highest degree, where r_u is the rate constraint of u. However, such a strategy would miss the path altogether. In contrast, an adversary with a global knowledge of the graph can plan ahead and infect the whole path, by starting from the root of the tree but spending its budget to infect the next node on the path (and the extra budget to infect part of tree). While greedy would infect only $O(1)$ nodes, a well-planned adversary would infect $n - O(1)$ nodes.

1.1 Our Results

We first show that the problem of computing an optimal and near-optimal adversarial strategies under typical constraints is computationally hard. We then identify necessary and sufficient conditions in terms of network structure and edge infection probabilities such that the adversarial process can infect polynomially more nodes than the stochastic process while maintaining a similar "signature" as the oblivious stochastic infection strategy. Our first set of results show that when the minimum weighted graph cut is $\Omega(\log n)$ (on a network with n nodes), the standard oblivious stochastic infection strategy can essentially infect all nodes. Thus the interesting regime to analyze is when the minimum weighted cut is $o(\log n)$. In this regime we demonstrate that the optimal adversarial infection can be pandemic (namely infect all nodes) while the oblivious stochastic strategy infects, in expectation, a constant number of nodes. We then identify a surprising connection between an additional structural quantity of interest in a network, the network *toughness*, and adversarial infections. Based on the network toughness, we characterize conditions guaranteeing the existence of a pandemic adversarial strategy as well as its efficient computation.

1.2 Related Work

Most work on computer malware has been focused on virus and worm propagation [17, 19]. The vast majority of the literature has focused on modelling and simulations of the behavior of a stochastic malware which spreads according to the Independent Cascades model or its extension to repeated infection attempts, the Susceptible-Infected model [7, 19–22]. However, none of these papers consider intelligently-designed malware that can choose which nodes to infect based on some prior computation or knowledge of the network.

Another line of research that we would like to mention is the one devoted to error and attack tolerance in online networks; see the seminal work of [1] and its long line of follow-up research such as [9]. The main thread of research is devoted to understanding how attacking and removing, in an unweighted graph, nodes of high-degree results in more parts of the network becoming disconnected than attacking and removing the same number of nodes, obliviously at random. In contrast to this work, we are interested in analyzing cascading effects that spread through the network, rather than a single attack and removal of nodes (and edges). Furthermore, our main interest is in coordinated attacks that are strategically designed, and so the type of node first targeted is chosen as to maximize a global effect in a provable way (resulting in infecting as many nodes as possible) rather than based on local heuristics (choosing high degree nodes).

1.3 Outline

In section 2 we provide a detailed definition of what comprises an adversarial infection and the description of the behavior of the oblivious stochastic infection strategy. In section 3 we discuss the computational complexity of computing an optimal, as well as near optimal, adversarial infection strategies. In section 4 we provide necessary conditions in terms of the network cuts such that the adversarial process can infect polynomially more nodes than the stochastic process. In section 5 we provide necessary and sufficient conditions for the existence of a pandemic adversarial infection strategy and provide efficient algorithms for computing such a strategy. Finally, in section 6 we summarize our contributions and list several intriguing directions for future research.

2 Model and Preliminaries

The Input Network. We are given an undirected, edge-weighted network $G = (V, E, p)$ on $|V| = n$ nodes, $|E| = m$ edges and an edge weight function $p : E \to (0, 1]$. We think of the weight $p(u, v)$ as the average amount of communication between neighbors u and v over a typical period of time. We will be particularly interested in the behavior of adversarial infections and the independent cascade model on the input network.

Adversarial Diffusion. An adversary strategy A is a rooted tree T_A, rooted at its seed node of choice (namely, the source of infection), that spans some arbitrary subset S of nodes. Each node u in T_A is responsible for infecting its children in T_A. For any subset X of nodes in T_A, let

$$\text{infect}(A, X) = \{v \in T_A \colon v\text{'s parent in } T_A \text{ belongs to } X\}.$$

We say that an adversary strategy A (with its rooted tree T_A) infects a set S of nodes, while obeying the first order constraints, if its rooted tree T_A spans S such that for every subset $X \subseteq S$ of nodes we have,

$$|\text{infect}(A, X)| \leq \left\lceil \sum_{u \in X} \sum_{v:(u,v)\in E} p(u, v) \right\rceil .$$

Namely, the number of node infections attributed to X is constrained by the ceiling of the total weight adjacent to the set X.

In this paper we only consider adversarial infections that obey the first-order constraints; unless stated explicitly otherwise, an adversary would be assumed to obey the first-order constraints. We will call the problem of finding a first-order constrained adversarial infection that maximizes number of infections *the adversarial infection problem* [3]. An adversarial strategy will be called *pandemic* if it infects all nodes.

Independent Cascades Diffusion Model. The Independent Cascades (IC) model of diffusion was formalized by Kempe et al. [13] and is by now a standard model

[3] It is not hard to see that the optimal adversary is deterministic: any stochastic adversary is a convex combination of trees so one can take the best one w.l.o.g.

to describe infection propagation in social and other contact networks [10]. The IC model can be thought of as a discrete version of the known *Susceptible-Infected-Removed (SIR) model.* The IC diffusion spreads via a random process, beginning at its seed node of choice from V. The process proceeds in rounds. In each round, a node u that got infected in the previous round gets a chance to subsequently infect each healthy neighbor v, with probability equal to the weight of the edge (u, v), namely $p(u, v)$. The node u becomes then recovered and does not spread the virus any further. If multiple nodes try to infect a new node in the current round, then each succeeds, independently, according to the corresponding edge weight. In this paper, we will focus on the the IC process only on undirected graphs. Throughout the text, we shall also refer to the stochastic infection strategy, as defined by the IC process, as the *oblivious stochastic strategy* or sometime as the *stochastic diffusion.*

Quantities of interest. All graphs considered in the paper are undirected graphs. Throughout our analysis we will frequently refer to the minimum weighted cut and maximum weighted cut in the weighted input graph: the minimum (resp. maximum) weighted cut, denoted $\Phi_G^{\min}$ (resp. $\Phi_G^{\max}$), is the value of the weighted cut $(S, \overline{S})$ that minimizes (resp. maximizes) the weighted sum

$$\Phi_G(S) := \sum_{\{u \in S, v \in \overline{S}, (u,v) \in E\}} p(u, v) \, .$$

When the identity of G is clear, we will often omit the subscript G.

We denote by $|S|$ the size of a set S. We denote the degree of a node u in a graph G by $d_G(u)$.

All logarithms in the paper are in base 2.

3 Hardness of Maximizing Adversarial Infection

In this section, we show that in general, the adversarial infection problem is hard to approximate under first order constraints.

Theorem 1. *The adversarial maximization problem is $2^{(\log^{1-\epsilon} n - 1)}$-hard to approximate, for any $\epsilon > 0$, unless $\boldsymbol{NP} \subseteq \boldsymbol{DTIME}(2^{O(\log^{1/\epsilon} n)})$.*

Proof. The proof is by reduction from the longest path problem on undirected graphs. Let $G(V, E)$ be the input graph for the longest path problem. We create an instance of the adversarial infection problem from G as follows. Starting with the graph G, we attach $(n - d_G(u))$ auxiliary vertices to each vertex $u \in V$. Assign a weight of $1/n$ to each edge in the resulting graph. Let H be the resulting graph.

Note that any infection strategy in H obeying the first-order constraints must be a path, since the total incident weight on any node in H is at most 1 (and is exactly 1 for nodes from G). If the path length is ℓ in H it translates to a path of length at least $\ell - 2$ in G (which is at least $\ell/2$ for $\ell \geq 4$), and a path of length ℓ in G translates to an infection strategy following a path of length ℓ

in H. The longest path problem is $2^{(\log^{1-\epsilon} n)}$-hard to approximate in undirected graphs, unless $\mathbf{NP} \subseteq \mathbf{DTIME}(2^{O(\log^{1/\epsilon} n)})$ [12], and the result easily follows.

We next show that the problem remains hard to approximate to within any constant factor even when the input instances are restricted to regular graphs with uniform infection probabilities.

Theorem 2. *The adversarial maximization problem does not admit a constant factor approximation in undirected regular graphs with uniform edge weights, unless* $\boldsymbol{P} = \boldsymbol{NP}$.

Proof. Let G be an undirected k-regular graph, $k \geq 2$ with uniform infection probability of $1/k$ on edges. Now the problem of finding a good strategy obeying first-order constraints becomes exactly the problem of finding a long path in the graph. Thus the problem of maximizing adversarial infection is as hard to approximate as the longest path on regular graphs. Even in 3-regular Hamiltonian graphs, the longest-path problem is known not to have any constant factor approximation, unless $\mathbf{P} = \mathbf{NP}$ [5].

4 A Cut-Based Analysis

We next proceed to exploring networks where an adversarial infection can infect polynomially more nodes than the oblivious stochastic strategy. We will show that two important parameters in understanding this goal is the size of the minimum weighted cut, $\Phi_G^{\min}$, and the size of maximum weighted cut, $\Phi_G^{\max}$, in the input graph G.

We first show that if $\Phi_G^{\min}$ is at least logarithmically large the oblivious stochastic strategy is essentially pandemic.

Theorem 3 (Theorem 1 of [2]). *Let* $G = (V, E, p)$. *For every positive constant* b, *there exists a constant* $c = c(b) > 0$ *so that if* $\Phi_G^{\min} \geq c\log(n)$, *then the probability that a realization of* G *is disconnected, where each edge* $(u, v) \in E$ *is kept with probably* $p(u, v)$, *is at most* $\frac{1}{n^b}$.

By the theorem we conclude,

Corollary 1. *For* c *large enough, the oblivious stochastic strategy would infect at least* $n - o(1)$ *nodes in expectation.*

Thus in this parameter regime, an improvement using adversarial infection strategies would be non-significant over the oblivious stochastic strategy. We note, however, that having a logarithmically large minimum cut is a quite stringent condition; in particular, any graph that has even one node with degree of size $o(\log(n))$ would violate the minimum cut condition. It is thus of high interest to analyze the regimes when this condition is violated. For this purpose we next show that no adversarial strategy can infect more than $\Theta(\Phi_G^{\max})$ nodes, and that when $\Phi_G^{\max}$ is large enough, a polynomial gap between the adversary to the oblivious stochastic strategy is feasible.

Theorem 4. *For any graph G no adversarial infection strategy obeying first-order constraint, as well as the oblivious stochastic strategy, can infect more than $\Theta(\Phi_G^{\max})$ nodes.*

On the other hand, for any value n there exists a graph G on n nodes with $\Phi_G^{\max} = \Theta(n)$ such that the optimal adversarial strategy obeying first order constraint infects $\Theta(\Phi_G^{\max})$ nodes while the oblivious stochastic strategy infects only $O(1)$ nodes in expectation. Moreover, the gap result holds on graphs where the edge infection probability is uniform and a constant (independent of n).

Proof. To prove the first part we make use of the following known fact that can be easily proven using the probabilistic method.

Fact 1 *For any weighted undirected graph $G = (V, E, p)$,*

$$\Phi_G^{\max} \geq 1/2 \sum_{(u,v)\in E} p(u, v) .$$

Now consider any adversarial infection strategy A, and let T_A be its tree of infections; let S be the set of nodes infected. Ignoring the root of T_A, either the total number of nodes at the odd levels of T_A must be at least $(|S| - 1)/2$ or the total number of nodes at the even levels of T_A must be at least $(|S| - 1)/2$. Since all infections at odd levels have to be attributed to nodes at the even levels (and vice versa), if the tree T_A conforms to first-order constraints, then we must have

$$(|S| - 1)/2 \leq \lceil 2\Phi_G^{\max} \rceil.$$

Since this holds for any choice of tree T_A (and hence any adversarial strategy), the claim follows for any adversary. A simple argument shows the claim for the oblivious stochastic process: any node u can infect at most $\sum_{v:(u,v)\in E} p(u, v)$ new nodes in expectation, and thus the total number of infections is, in expectation, at most

$$\sum_{u\in V} \sum_{v:(u,v)\in E} p(u, v) = \Theta(\Phi_G^{\max}).$$

We now prove the other part of the theorem. To show this we need to provide a graph G on n nodes such that the optimal adversary can infect $\Phi_G^{\max} = \Theta(n)$ nodes while the oblivious stochastic strategy infects $O(1)$ in expectation.

For simplicity of exposition assume that n is even. Take a cycle on $n/2$ nodes and set each edge probability on the cycle to be $1/2$. Now attach to each node u_i on the cycle, an auxiliary nodes v_i using also an edge of probability $1/2$. Note that by fact 1, the maximum weighted cut $\Phi_G^{\max}$ is $\Omega(n)$. Clearly, $\Phi_G^{\max} \leq |E| = n+1$ and so $\Phi_G^{\max} = \Theta(n)$, as required.

The adversary chooses all n nodes on the cycle (infection tree is a path), and no auxiliary nodes. This satisfies first-order constraints because each node u_i has an infection budget of at least 1 and it needs to infect exactly one node.

However, for any choice of the seed vertex, a stochastic strategy obtains only $O(1)$ nodes in expectation. To see this note that the infection survives for k steps on the cycle with probability at most $2/2^k$ and so it can infect, in expectation, at most $\sum_k 8k/2^k = 16 = O(1)$ nodes.

5 Pandemic Infections

In this section, we further explore the setting where the size of the minimum expected cut is $o(\log n)$. As we have seen earlier, in this setting the gap between an adversarial infection (obeying first-order constraints) and oblivious stochastic diffusion can be as large as $\Omega(n)$. We obtain here sufficient and necessary conditions for an adversarial infection to become pandemic (i.e., infect all nodes) by relating existence of such strategies to the notion of *toughness* of the graph.

The notion of graph toughness was first introduced in order to study conditions for the existence of Hamiltonian cycles in graphs, see [4]. Given an undirected graph $G = (V, E)$ and a subset of nodes S, let $|S|$ be the size of S and $c_G(S)$ be the number of connected components in the graph induced on $V \setminus S$ obtained from G by deleting all nodes in the set S. The toughness of the graph G, where G is not the complete graph, denoted by $\tau(G)$ is defined as follows

$$\tau(G) = \min_{S \subset V,\, c_G(S)>1} \frac{|S|}{c_G(S)} . \tag{1}$$

The toughness of the complete graph is defined to be infinity. Toughness of a cycle, on the other hand, is 1 since by deleting any subset of k nodes, we can create at most k connected components, and removing two nodes with no edge between creates exactly two components. As another example, it is easy to verify that the toughness of a tree with maximum degree Δ (and at least three nodes) is $1/\Delta$. It is also easy to verify that the toughness of a graph is positive if and only if the graph is connected.

There is a vast literature on the connections between graph toughness and spanning trees; for a recent survey see [15]. In what follows, we will show a close connection between existence and algorithms for adversarial infections that are pandemic to the toughness of the underlying connections graph.

We start by developing sufficient conditions under which there exists a spanning tree that can be exploited by an adversary obeying the first-order constraints.

Theorem 5. *For any connected, weighted undirected graph $G = (V, E, p)$ such that*

$$\forall u \in V, \sum_{v:(u,v)\in E} p(u, v) \geq \left\lceil \frac{1}{\tau(G)} + 1 \right\rceil \tag{2}$$

there is an adversary strategy that infects all nodes in G and obeys the first-order constraints. Moreover, assume that

$$\forall u \in V, \sum_{v:(u,v)\in E} p(u, v) \geq \left\lceil \frac{1}{\tau(G)} + 2 \right\rceil . \tag{3}$$

Then, there is a polynomial-time computable *adversary strategy that infects all nodes in G and obeys the first-order constraints.*

Proof. By Win's Theorem [18], every undirected graph G with toughness $\tau(G)$ has a spanning tree with maximum degree bounded by $d = \lceil(1/\tau(G)+2)\rceil$. Let T be such a tree. Root this tree at a leaf node s and let this node be the seed. The adversary strategy is to infect all nodes of G, starting with node s, by using the edges of T. Now no node is responsible for infecting more than $d-1$ children and so the first-order constraints are obeyed: for each node u, its degree minus one is at most $\lceil(1/\tau(G)+2)\rceil - 1 = \lceil(1/\tau(G)+1)\rceil \leq \sum_{v:(u,v)\in E} p(u,v)$. To show the other part of the theorem we make use of an algorithmic result by Fürer and Raghavachari [11] that states that if there exists a spanning tree of degree at most Δ than one can construct in polynomial time a spanning tree of degree at most $\Delta + 1$. The assertion of the theorem then follows similarly to the previous case, where now for each node u, $\lceil 1/\tau(G) + 2\rceil \leq \sum_{v:(u,v)\in E} p(u,v)$.

Example 1. To illustrate the theorem, consider the following example. Take a random graph $G \in \mathrm{G}(n,p)$ with $p \geq 2\log n/n$. With probability of $1 - o(1)$, G contains a Hamiltonian cycle (see for instance [14]). So $\tau(G) \geq 1$, and thus to apply Theorem 5, all we need is that for every node u, $\sum_{v:(u,v)\in E} p(u,v)$ is at least 2. By the Chernoff bound, all vertex degrees in G are at least $pn/2$ with high probability. Hence by setting $p(u,v) = 4/pn$ on each edge (u,v), we satisfy the conditions of Theorem 5, and can conclude existence of an adversarial strategy that leads to pandemic infection. Note that as network degrees increase in this example, smaller and smaller edge infection probabilities suffice to get adversarial pandemic infection.

Discussion:

We note that if we slightly weaken the condition stated in Theorem 5, the result no longer holds. Specifically, if we allow just $O(1)$ nodes to violate condition (2) by only a slack of 1, that is, allow $O(1)$ nodes u with $\sum_{u:(u,v)\in E} p(u,v) = \lceil(1/\tau(G)\rceil$, then for infinitely many toughness values there will exist infinitely many networks such that the optimal adversarial strategy is not pandemic (namely, infects all nodes). To see this, let $k \geq 2$ be an integer and consider the following family of graphs $H_{n,k}$, where $n \geq 3k(k+3)+3$ is integral. Take three nodes v_1, v_2, v_3 as connect them as a clique. In addition create $3k$ vertex-isolated cliques, indexed $C(i,j)$ for $1 \leq i \leq 3, 1 \leq j \leq k$, each on $(n-3)/3k$ nodes. To complete the construction connect each node v_i to some node in each clique $C(i,j)$ ("a representative"), where $1 \leq j \leq k$. Thus the degree of each node v_i in the construction is $k+2$, in addition to being connected to k representatives it is also connected to the other v_i nodes. Set all edge weights to 1, except for each i the edges connecting v_i to the cliques $C(i,1)$ and $C(i,2)$; the weight on each of these edges is set to 1/2.

We now observe a few simple facts about the graph $H_{n,k}$. First, its toughness is $1/(k+1)$. Second, it has a spanning tree (tree spanning all graph nodes). Third, in each of its spanning trees each v_i must be connected to a representative from each $C(i,j)$ for $1 \leq j \leq k$. In addition, one of the v_is must be connected to the two other v_is (otherwise the spanning tree is disconnected), and so its degree in the spanning tree must be $k+2$. Last, except for the three v_is, the sum of edge-weights touching any node is at least $k+2$ (as $\frac{n-3}{3k} \geq k+3$ condition (2) holds for all clique nodes). However, the sum of edge-weights touching a v_i, namely $\sum_{u:(v_i,u)\in E} p(v_i,u)$, is $2+k-2+1/2\cdot 2 = k+1$. Thus the degree of each of the v_is in a tree representing an adversarial infection can be at most $k+1$, which is strictly smaller than $k+2$, and so any such adversary cannot infect all nodes in the graph. In fact, a constant fraction of the graph nodes will not be infected — all the nodes belonging to one of the cliques $C(i,j)$.

□

We now show that for a given value of the toughness τ, there exist infinitely many graphs such that the stochastic diffusion can only infect $O(\tau)$ nodes, for $\tau \geq 4$. In light of theorem 5, on such graphs the gap between the stochastic diffusion and the adversarial diffusion is large.

Theorem 6. *For any value of the toughness $\tau \geq 4$ and positive integer $\ell \geq 3$, there exists a weighted undirected graph $G = (V, E, p)$ on $n = \tau\ell$ nodes and toughness τ, such that condition (2) is satisfied (hence the adversarial process can infect all n nodes), yet the stochastic diffusion infects only $O(\tau)$ nodes in expectation.*

Proof. Take a cycle with ℓ nodes, say, $v_0, v_1, ..., v_{\ell-1}$. Now replace each v_i by a clique C_i on $\tau \geq 4$ nodes. Now for each i, connect vertices in clique C_i to vertices in C_{i+1} by a complete bipartite graph. Note that the total number of nodes n equals $\tau\ell$. Also, one can verify that the toughness of this modified cycle is τ, since the toughness of a simple cycle is 1.

Assign a probability of 1 to edges inside each C_i, and probability $\frac{1}{2\tau^2}$ to edges between the cliques. Note that the probability that any edge between two adjacent cliques gets realized is less than 1/2. Indeed, the probability that no edge between two adjacent cliques gets realized is

$$(1 - 1/(2\tau^2))^\tau \geq 1 - \frac{1}{2\tau} > 1/2\,,$$

where we used the inequality $(1+x)^r \geq 1 + rx$, for $x \geq -1$ and $r \in \mathbb{R} \setminus (0,1)$. In particular, the behavior of the stochastic diffusion on this network is essentially as it is on a cycle with edge probability less than 1/2 that was analyzed in the proof of Theorem 4; the only difference is that now each node on the cycle infects as well all the nodes in its clique. Thus the stochastic diffusion infects at most $O(\tau)$ nodes in expectation. Finally, condition (2) trivially holds since $\tau \geq 4$ and so each vertex has a probability mass of at least 3 incident on it, while $1/\tau \leq 1$.

6 Conclusions and Future Work

In this work we initiated the study of *adversarial* infection strategies which can decide intelligently which nodes to infect next in order to maximize their spread, while obeying first-order constraints as to not get discovered. We have demonstrated that a well-planned adversarial infection can substantially increase the number of nodes infected with respect to the standard Independent Cascades infection strategy. We designed necessary and sufficient conditions to understand when this is possible. Based on novel connection to the network toughness, we characterize networks where existence of adversarial strategies that are pandemic (infect all nodes) is guaranteed, as well as efficiently computable.

Our results have focused on first order constraints: keeping the traffic involving any set X lower than the ceiling of its expected value, namely,

$$\left\lceil \sum_{u \in X} \sum_{v:(u,v) \in E} p(u,v) \right\rceil .$$

An interesting future direction is to consider flow constraints where the number of infections caused by any set X is at most the flow leaving X, namely,

$$\left\lceil \sum_{u \in X} \sum_{v \notin X:(u,v) \in E} p(u,v) \right\rceil .$$

Moreover, it would be interesting to consider directed graphs as well.

Another interesting avenue for further exploration is to analyze the effectiveness of vaccination strategies designed for controlling stochastic epidemics in limiting the spreading of the adversarial epidemic. In particular, one could consider immunization strategies such as immunizing high degree nodes or acquaintance immunization based on the immunization of a small fraction of random neighbors of randomly selected nodes. Such strategies are known to be effective at controlling stochastic epidemics [6, 8] but might be ineffective containing the first-order constrained adversarial infection. To demonstrate this, let us go back to the example of the introduction, namely a path connected at one of its ends to the root of a two-level binary tree. Consider vaccinating the node with the highest weighted degree, which is in this example the root of tree. Vaccinating the root of the tree will not help in containing adversarial infections, which would still infect $n - O(1)$ nodes (the path nodes). Yet vaccinating the middle node of the path is much better for containment — yielding at most $n/2 + O(1)$ infected nodes by an adversarial epidemic. It is therefore interesting to understand and develop vaccination schemes that aim to minimize the number of infections under the adversarial setting.

7 Acknowledgements

We would like to thank the anonymous reviewers for their helpful comments.

Moez Draief was supported by QNRF grant NPRP-09-1150-2-448.

Sanjeev Khanna was supported in part by the National Science Foundation grants CCF-1116961 and IIS-0904314.

References

1. Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

2. Noga Alon. A note on network reliability. In *Discrete Probability and Algorithms*, pages 11–14. Springer, 1995.
3. Justin Balthrop, Stephanie Forrest, M. E. J. Newman, and Matthew M. Williamson. Technological networks and the spread of computer viruses. *Science*, 304(5670):527–529, 2004.
4. Douglas Bauer, Hajo Broersma, and Edward F. Schmeichel. Toughness in graphs - a survey. *Graphs and Combinatorics*, 22(1):1–35, 2006.
5. Cristina Bazgan, Miklos Santha, and Zsolt Tuza. On the approximation of finding a(nother) Hamiltonian cycle in cubic Hamiltonian graphs. *J. Algorithms*, 31(1):249–268, 1999.
6. Tom Britton, Svante Janson, and Anders Martin-Löf. Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Advances in Applied Probability*, 39(4):922–948, 2007.
7. Zesheng Chen, Chao Chen, and Chuanyi Ji. Understanding localized-scanning worms. In *IPCCC*, pages 186–193, 2007.
8. Reuven Cohen, Shlomo Havlin, and Daniel Ben-Avraham. Efficiency immunization strategies for computer networks and populations. *Physical Review Letters*, 91(24):247901–1—247901–4, 2003.
9. Paolo Crucitti, Vito Latora, Massimo Marchiori, and Andrea Rapisarda. Efficiency of scale-free networks: Error and attack tolerance. *Physica A*, 320(642):622–642, 2003.
10. David A. Easley and Jon M. Kleinberg. *Networks, Crowds, and Markets - Reasoning About a Highly Connected World.* Cambridge University Press, 2010.
11. Martin Fürer and Balaji Raghavachari. Approximating the minimum-degree Steiner tree to within one of optimal. *J. Algorithms*, 17(3):409–423, 1994.
12. David R. Karger, Rajeev Motwani, and G. D. S. Ramkumar. On approximating the longest path in a graph. *Algorithmica*, 18(1):82–98, 1997.
13. David Kempe, Jon M. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *KDD*, pages 137–146, 2003.
14. J. Komlós and E. Szemerédi. Limit distributions for the existence of Hamiltonian circuits in a random graph. *Discrete Mathematics*, (43):55–63, 1983.
15. Kenta Ozeki and Tomoki Yamashita. Spanning trees: A survey. *Graphs and Combinatorics*, 27(1):1–26, January 2011.
16. Giuseppe Serazzi and Stefano Zanero. Computer virus propagation models. In *MASCOTS Tutorials*, pages 26–50, 2003.
17. Nicholas Weaver, Vern Paxson, Stuart Staniford, and Robert Cunningham. A taxonomy of computer worms. In *WORM*, pages 11–18, 2003.
18. Sein Win. On a connection between the existence of k-trees and the toughness of a graph. *Graphs and Combinatorics*, 5(1):201–205, 1989.
19. Guanhua Yan, Guanling Chen, Stephan Eidenbenz, and Nan Li. Malware propagation in online social networks: nature, dynamics, and defense implications. In *ASIACCS*, pages 196–206, 2011.
20. Cliff Changchun Zou, Donald F. Towsley, and Weibo Gong. Email worms modeling and defense. In *ICCCN*, pages 409–414, 2004.
21. Cliff Changchun Zou, Donald F. Towsley, and Weibo Gong. On the performance of internet worm scanning strategies. *Perform. Eval.*, 63(7):700–723, 2006.
22. Cliff Changchun Zou, Donald F. Towsley, Weibo Gong, and Songlin Cai. Routing worm: A fast, selective attack worm based on IP address information. In *PADS*, pages 199–206, 2005.