An Efficient Curing Policy for Epidemics on Graphs

Kimon Drakopoulos, *Member, IEEE*, Asuman Ozdaglar, and John N. Tsitsiklis, *Fellow, IEEE*

Abstract—We provide a dynamic policy for the rapid containment of a contagion process modeled as an SIS epidemic on a bounded degree undirected graph with n nodes. We show that if the budget r of curing resources available at each time is $\Omega(W)$, where W is the CutWidth of the graph, and also of order $\Omega(\log n)$, then the expected time until the extinction of the epidemic is of order $O(n/r)$, which is within a constant factor from optimal, as well as sublinear in the number of nodes. Furthermore, if the CutWidth increases only sublinearly with n, a sublinear expected time to extinction is possible with a sublinearly increasing budget r.

Index Terms—Networks, epidemics, control, contagion, influence minimization

1 Introduction

Many contagion processes over large networks can lead to costly cascades unless controlled by outside intervention. Examples include epidemics spreading over a population of individuals, viruses attacking a network of connected computers, or financial contagion in a network of banks. In this paper we study how this type of contagion can be prevented or contained by dynamically curing some of the infected nodes under a budget constraint on the amount of curing resources (curing budget) that can be deployed at each time.

More specifically, we consider a canonical SIS epidemic model on an undirected graph[1] with n nodes, with a common infection rate along any edge that connects an infected and a healthy node, and node-specific curing rates $\rho_v(t)$ at each node v. The curing rates are to be chosen according to a curing policy which is based on the past history of the process and the network structure, subject to an upper bound on the total curing rate $\sum_v \rho_v(t)$ at every time instant t. Curing policies are evaluated in terms of the expected time it takes for the epidemic to become extinct, i.e., the process to reach the state where all nodes are healthy.

The main contribution of this paper is the construction of a policy for the dynamic control of epidemics. Our analysis involves the CutWidth of the underlying graph. Intuitively, the CutWidth measures the required budget of curing resources in a simpler deterministic curing problem, in which infected nodes are cured one at a time, subject to the constraint that the number of edges between healthy and infected nodes is at all times less than or equal to the budget of curing resources. Our policy, called from now on the CURE policy, possesses several desirable properties:

i) Assuming that the available curing resources are larger than a certain quantity that depends on several global characteristics of the underlying network (maximum degree and CutWidth) and considering the worst case where all nodes are initially infected, our policy is (order) optimal.

ii) When a strict subset of the nodes is initially infected our policy is (order) optimal with high probability if the available curing resources are larger than a certain quantity that depends on local properties of the set of initially infected nodes ("impedance") and the maximum degree of the graph.

iii) In a companion paper [6], we show that for certain bounded degree graphs a contagion process cannot be rapidly contained, i.e., the expected time to extinction cannot be made sublinear in the number n of nodes using a sublinear curing budget at each time. Specifically we establish that if the CutWidth increases at least linearly with n, then a sublinear (in n) expected time to extinction is impossible with a sublinear budget r. In this paper we prove that for graphs with bounded degree and sublinear CutWidth, a sublinear curing budget at each time is enough to guarantee sublinear extinction time. The combination of these two results provides a qualitative characterization of the best possible scaling of the extinction time in terms of the CutWidth.

Our policy is based on a combinatorial result which states the following. Given an initial set of infected nodes, nodes can be removed from that set, one at a time, in way that the maximum cut (number of edges) between healthy and infected nodes encountered during this process is upper bounded by the sum of the CutWidth of the graph and the cut associated with the initial set. Let us refer to the sequence of subsets encountered during this process as a *target path*. The main idea underlying our policy is to allocate the entire curing budget to appropriate nodes so that we stay most of

1. Our results actually are easily generalized to the case of directed graphs.

• *The authors are with the Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139. E-mail: {kimondr, asuman, jnt}@mit.edu.*

Manuscript received 24 June 2014; revised 9 Nov. 2014; accepted 31 Dec. 2014. Date of publication 14 Jan. 2015; date of current version 11 Feb. 2015. Recommended for acceptance by R. Sepulchre.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TNSE.2015.2393291

the time, with high probability, on or near the target path. We show that this is indeed possible, as long as the curing budget scales in proportion to the CutWidth. We also show that the policy is optimal (within a multiplicative constant) if the available budget at each time is also $\Omega(\log n)$.[2]

A similar model, but in which the curing rate allocation is done statically (open-loop) has been studied in [3], [5], [8], [14], but the proposed methods were either heuristic or based on mean-field approximations of the evolution process. Closer to our work, the authors of [2] let the curing rates be proportional to the degree of each node, but independent of the current state of the network, which means that curing resources may be wasted on healthy nodes. On a graph with bounded degree, the policy in [2] achieves sublinear time to extinction, but requires a curing budget that is proportional to the number of nodes. In contrast, our policy achieves the same performance (sublinear time to extinction) for all bounded degree graphs with small CutWidth by properly focusing the curing resources. As an extreme example, consider a line graph with n nodes, and assume that the $n/2$ leftmost nodes are initially infected. The degree-based policy of [2] requires a total budget proportional to n and allocates it proportional to the degree. In contrast, our policy can achieve sublinear expected time to extinction with a $\Omega(\log n)$ but sublinear budget. This is because, instead of allocating the available budget to all nodes, our policy focuses on specific nodes on the boundary between healthy and infected nodes, in this instance on the rightmost infected node. By extending this idea, our policy achieves a similar improvement for all graphs with sublinear CutWidth. More specifically, for a graph with maximum degree Δ the degree based policy of [2] achieves expected extinction time which is $O(\log n)$ when the curing budget at each time is at least $\Omega(n\Delta)$. In contrast, for a graph with maximum degree Δ and CutWidth at most $O(\log n/n)$, our policy achieves $O(\log n)$ expected extinction time when the curing budget per time instant is at least $\Omega(n/log\, n)$.

The rest of the paper is organized as follows. In Section 2 we present the details of the model. In Section 3 we introduce the CutWidth and establish the combinatorial result mentioned earlier. In Section 4 we present the policy and analyze its performance. In Section 5 we develop some corollaries that demonstrate the possibility of fast extinction using a sublinear budget and the approximate optimality of our policy in a certain regime. We also mention some examples. Finally, in Section 7 we offer some closing remarks.

2 The Model

We consider a network, represented by a connected undirected graph $G = (V, E)$, where V denotes the set of nodes and E denotes the set of edges. We use n to denote the number of nodes. Two nodes $u, v \in V$ are *neighbors* if $(u, v) \in E$. We denote by Δ the maximum of the node degrees.

We assume that the nodes in a set $I_0 \subseteq V$ are initially infected and that the infection spreads according to a controlled contact process where the rate at which infected nodes get cured is determined by a network controller. Specifically, each node can be in one of two states: *infected* or *healthy*. The controlled contact process—also known as the SIS epidemic model—on G is a right-continuous, continuous-time Markov process $\{I_t\}_{t\geq 0}$ on the state space $\{0,1\}^V$, where I_t stands for the set of infected nodes at time t. We refer to I_t as the *infection process*.

State transitions at each node occur independently according to the following dynamics.

a) The process is initialized at the given initial state I_0.
b) If a node v is healthy, i.e., if $v \notin I_t$, the transition rate associated with a change of the state of that node to being infected is equal to an infection rate β times the number of infected neighbors of v, that is,

$$\beta \cdot \big|\{(u, v) \in E : u \in I_t\}\big|,$$

where we use $|\cdot|$ to denote the cardinality of a set. By rescaling time, we can and will assume throughout the paper that $\beta = 1$.
c) If a node v is infected, i.e., if $v \in I_t$, the transition rate associated with a change of the state of that node to being healthy is equal to a curing rate $\rho_v(t)$ that is determined by the network controller, as a function of the current and past states of the process. We are assuming here that the network controller has access to the entire past evolution of the process.

We assume a *budget constraint* of the form

$$\sum_{v\in V} \rho_v(t) \leq r, \tag{1}$$

for each time instant t, reflecting the fact that curing is costly. A *curing policy* is a mapping which at any time t maps the past history of the process to a curing vector $\rho(t) = \{\rho_v(t)\}_{v\in V}$ that satisfies (1).

We define the *time to extinction* as the time until the process reaches the absorbing state where all nodes are healthy:

$$\tau = \min\{t \geq 0 : I_t = \emptyset\}.$$

The *expected time to extinction* (the expected value of τ) is the performance measure that we will be focusing on.

3 Graph Theoretic Preliminaries

In this section we introduce the notions of a cut and of the CutWidth that will be used in the description of our policy. We state some of their properties and then proceed to develop a key combinatorial result that will play a critical role in the analysis of our policy's performance. Throughout, we assume that we are dealing with a particular given graph G.

3.1 CutWidth

For convenience, we will be using the shorthand term "bag" to refer to "a subset of V." We also use the following notation. For any two bags A and B, and any $v \in V$, we let

$$A \setminus B = \{v \in A : v \notin B\},$$

and

$$A - v = A \setminus \{v\}.$$

We also use A^c to denote the complement, $V \setminus A$ of A.

2. We write $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. We write $f(n) = \Omega(g(n))$ if $\liminf_{n\to\infty} f(n)/g(n) > 0$. Finally, we write $f(n) = O(g(n))$ if $\limsup_{n\to\infty} f(n)/g(n) < \infty$.

We next define the concept of a *monotone crusade*. A monotone crusade from a bag A to another bag B, where $B \subseteq A$, is a finite sequence of bags that starts with A and ends with B, so that at each step of the sequence no nodes are added (cf. Part (iii) of Definition 1), and exactly one node is removed (cf. Part (iv) of Definition 1).

Definition 1. *For any two bags A and B, with $B \subseteq A$, a (monotone) crusade from A to B, or $(A \downarrow B)$-**crusade** for short, is a sequence $\omega = (\omega_0, \omega_1, \ldots, \omega_k)$ of bags of length $|\omega| = k+1$, with the following properties:*

i) $\omega_0 = A$,
ii) $\omega_k = B$,
iii) $\omega_{i+1} \subset \omega_i$, *for* $i = 0, 1, \ldots, k-1$, *and*
iv) $|\omega_i \setminus \omega_{i+1}| = 1$, *for* $i = 0, 1, \ldots, k-1$.

We denote by $\mathcal{C}(A \downarrow B)$ the set of all $(A \downarrow B)$-crusades.

The number of edges connecting a bag A with its complement is called the cut of the bag. It is equal to the total rate at which new infections occur, when the set of currently infected nodes is A.

Definition 2. *For any bag A, its **cut**, $\mathrm{c}(A)$, is defined as the cardinality of the set of edges*

$$\{(u,v) : u \in A,\ v \in A^c\}.$$

In Proposition 1 below, we record, without proof, an elementary property of cuts.

Proposition 1. *For any two bags A and B, we have*

$$\mathrm{c}(A \cup B) \leq \mathrm{c}(A) + \mathrm{c}(B) \leq \mathrm{c}(A) + \Delta \cdot |B|.$$

We define the width of a monotone crusade ω as the maximum cut encountered during the crusade. Intuitively, this is the largest infection rate to be encountered if the nodes were to be cured according to the sequence prescribed by the crusade deterministically, if no new infections happen in between.

Definition 3. *Given an $(A \downarrow B)$-crusade $\omega = (\omega_0, \ldots, \omega_k)$, its **width** $z(\omega)$ is defined by*

$$z(\omega) = \max_{0 \leq i \leq k} \{\mathrm{c}(\omega_i)\}.$$

We next define what we call the impedance of a bag A, as the minimum possible width among the $\mathcal{C}(A \downarrow \emptyset)$-crusades. This minimization captures the objective of finding a crusade along which the total infection rate is always small.

Definition 4. *The **impedance** $\delta(A)$ of a bag A is defined by*

$$\delta(A) = \min_{\omega \in \mathcal{C}(A \downarrow \emptyset)} z(\omega). \tag{2}$$

*For the special case where $A = V$, the impedance is known as the **CutWidth** [1], [4], [9], [12], and will be denoted by W. Fig. 1 illustrates examples of graphs with typical values of CutWidth.*

We say that a (monotone) crusade $(A \downarrow B)$-crusade ω is *optimal* if it attains the minimum in Eq. (2). It can be seen that the impedances satisfy the Bellman equation:

$$\delta(A) = \max\{\mathrm{c}(A), \min\{\delta(B) : B \subseteq A,\ |A \backslash B| = 1\}\}. \tag{3}$$

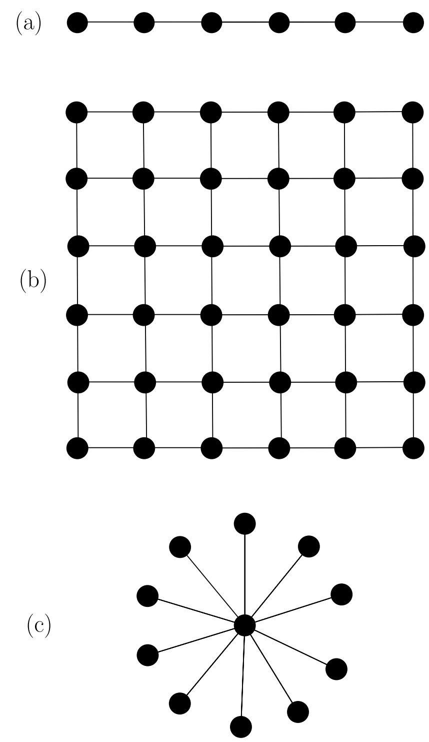

Fig. 1. (a) A line graph has CutWidth equal to 1. A corresponding monotone crusade starts with a bag containing all nodes and the subsequent bags are constructed by removing nodes one at a time from left to right. (b) A two dimensional $\sqrt{n} \times \sqrt{n}$ grid graph has CutWidth $\sqrt{n}+1$. A corresponding monotone crusade starts with a bag containing all nodes and the subsequent bags are constructed by removing nodes one at a time from left to right and row by row. (c) A star graph with $2n$ leaves has CutWidth equal to n. A corresponding monotone crusade starts with a bag containing all nodes and the subsequent bags are constructed by removing the first n leaves one at a time, then the center and then the remaining n. Expander graphs are another example of graphs with linear (in the size of the graph) CutWidth but with bounded degree.

Furthermore, along an optimal crusade, we have $\delta(\omega_{i+1}) \leq \delta(\omega_i)$, for $i = 0, 1, \ldots, k-1$. Finally, we note that $\mathrm{c}(A) \leq \delta(A)$.

3.2 Impedance and CutWidth

n this section we discuss the relation between the impedance of an arbitrary bag and the CutWidth. The impedance of a bag A is at least $\mathrm{c}(A)$, which in general may be much larger than the CutWidth.[3] This is a concern because the stochastic nature of the infections can always bring the process to a bag with high impedance, and therefore high subsequent infection rates. The next lemma provides an upper bound on the impedance of a bag A in terms of the CutWidth W of the graph and the cut of A.

Lemma 1. *For any bag A, we have*

$$\delta(A) \leq W + \mathrm{c}(A).$$

Proof. IConsider a monotone crusade $\omega \in \mathcal{C}(V \downarrow \emptyset)$ whose width is equal to the CutWidth W. This crusade starts with V and removes nodes one at a time, until the empty

3. As an example, consider a line graph, and let A be the set of even-numbered nodes. Then, $\mathrm{c}(A)$ is approximately n, whereas the CutWidth of the line graph is equal to 1.

set is obtained. Let $v_1, v_2, \ldots, v_n$ be the nodes in V, arranged in the order in which they are removed.

Let us now fix a bag A. We construct a monotone crusade $\hat{\omega} \in \mathcal{C}(A \downarrow \emptyset)$ as follows. We start with A and remove its nodes one at a time, according to the order prescribed by ω. For example, if $n = 4$, and $A = \{v_2, v_4\}$, the monotone crusade that starts from A first removes node v_2 and then removes node v_4.

At any intermediate step during the crusade $\hat{\omega}$, the current bag is of the form $A \cap \{v_k, \ldots, v_n\}$, for some k. It only remains to show that the cut of such a bag is upper bounded by $\mathrm{c}(A) + W$. Let $R = \{v_1, \ldots, v_{k-1}\}$. Note that

$$\mathrm{c}(R) \le W,$$

because of the definition of the width and the assumption that the width of ω is W. Note also that the current bag is simply $A \cap R^c$.

For any two sets S_1 and S_2, let $e(S_1, S_2)$ be the number of edges that join them. We have that

$$\begin{aligned}
\mathrm{c}(A \cap R^c) &= e\big(A \cap R^c, (A \cap R^c)^c\big) \\
&= e(A \cap R^c, A^c \cup R) \\
&\le e(A \cap R^c, A^c) + e(A \cap R^c, R) \\
&\le e(A, A^c) + e(R^c, R) \\
&= \mathrm{c}(A) + \mathrm{c}(R) \\
&\le \mathrm{c}(A) + W.
\end{aligned}$$

We conclude that the cut associated with any intermediate bag in the crusade $\hat{\omega}$ is upper bounded by $\mathrm{c}(A) + W$. It follows that the width of $\hat{\omega}$, and therefore $\delta(A)$ as well, is also upper bounded by that same quantity. □

4 THE CURE POLICY

In this section, we present our curing policy and we study the resulting expected time to extinction, starting from an arbitrary initial set of infected modes. Loosely speaking, the policy, at any time, tries to follow a certain desirable (monotone) crusade, called a target path, by allocating all of the curing resources to a single node, namely, the node that should be removed in order to obtain the next bag along the target path. On the other hand, this ideal scenario may be interrupted by infections, at which point the policy shifts its attention to newly infected nodes, and attempts to return to a bag on the target path. It turns out that under certain assumptions, this is successful with high probability and does not take too much time. However, with small probability, the process veers far off from the target path; in that case the policy "restarts" in a manner that we will make precise in the sequel.

It is quite intuitive (and formally established in [6]) that a fast (sublinear) time to extinction may not be possible if the curing budget is smaller than the CutWidth. For this reason, we focus on the regime where the curing rate is at least proportional to the CutWidth, and more concretely, on the regime where $r \ge 4W$, which we henceforth assume.

Under the above assumptions on the budget r, and the additional assumptions that $r = \Omega(\log n)$ and $r \ge 8\Delta$, we will construct a policy whose expected time to extinction is $O(n/r)$; cf. Theorem 1 and Corollary 1.

4.1 Description of the CuRe Policy

Waiting period. A typical attempt starts at some bag A, with a waiting period. (If this is the first attempt, then $A = I_0$. Otherwise, A is the bag at the end of the preceding attempt.) During the waiting period, all curing rates $\rho_v(t)$ are kept at zero.[4] The waiting period ends at the first subsequent time that[5]

$$\mathrm{c}(I_t) \le r/8.$$

Let B be the bag I_t right at the end of the waiting period, and let $\omega^B = (\omega_0^B, \ldots, \omega_{|B|}^B)$ the corresponding optimal crusade, which we refer to as the *target path*.

Segments. Each segment of an attempt starts either at the end of the waiting period or at the end of a preceding segment of the same attempt. In all cases, the segment starts with a bag on the target path. For the first segment, this is guaranteed by the definition of the target path. For subsequent segments, it will be guaranteed by our specifications of what happens at the end of the preceding segment. Let $v_1, \ldots, v_m$ be the nodes in the bag at the beginning of a segment, arranged in the order according to which they are to be removed along the target path. For example, the bag at the beginning of the segment is $\omega_0^B = \{v_1, \ldots, v_m\}$, the next bag is $\omega_1^B = \{v_2, \ldots, v_m\}$, etc. The node v_1 is called the *target node*; the goal of the segment is to cure the target node and reach the bag $C = \{v_2, \ldots, v_m\}$. For all t during the segment, we define $D_t = I_t \setminus C$; this is the set of infected nodes that do not belong to the next bag on the target path. At the beginning of the segment, $I_t = C \cup \{v\}$ and therefore $D_t = \{v_1\}$. During the segment, the entire curing budget is allocated to an arbitrarily chosen node from D_t. Note that $\rho_v(t) = 0$ for $v \in C$ during the segment and therefore, we always have $I_t \supseteq C$.

The segment ends when either:

i) all nodes have been cured, i.e., $I_t = \emptyset$; in this case, the attempt is considered successful and the process is over.

ii) $I_t = C$ and $C \ne \emptyset$ in which case the target node is cured, the process is on the target path, and we are ready to start the next segment. In this case, we say that we have a *short* segment.

ii) $|D_t| \ge r/8\Delta$, in which case we say that the segment was *long*, and that the attempt has failed. In this case, the attempt has no more segments, and a new attempt will be initiated, starting with a waiting period.

4.2 Performance Analysis—Outline

We now proceed to establish an upper bound on the expected time to extinction, under the assumption that

4. During the waiting period the curing budget is wasted and not allocated to any of the nodes. Note that the cut of I_t during the waiting phase could be linear in the number of nodes, while we focus on the regime where the available budget is sublinear. Therefore, regardless of the allocation, during the waiting period the process would have an upward drift. For this reason, allocating budget to a subset of nodes in this period would not have a significant effect on the performance.

5. Note that the waiting period is guaranteed to terminate in finite time, with probability 1. This is because if it were infinite, then healthy nodes would keep getting infected until eventually $I_t = V$. But $\mathrm{c}(V) = 0$, which means that at some point the condition $\mathrm{c}(I_t) \le r/8$ would be satisfied and the waiting period would be finite, a contradiction.

$r \geq 4W$, for any set of initially infected nodes. If the process always stayed on the target path, that is, if we had no infections, the expected time to extinction would be the time until all nodes (at most n of them) were cured. Given that nodes are cured at a rate of r, the expected time to extinction would have been $O(n/r)$. On the other hand, infections do delay the curing process, by increasing $|D_t|$ during segments, and we need to show that these do not have a major impact.

There are two kinds of segments to consider, *short* ones, at the end of which $|D_t| = 0$, and *long* ones, at the end of which $|D_t| \geq r/8\Delta$. During a segment, the size of D_t (the "distance" from the target path) is at most $r/8\Delta$. Using also an upper bound on the size of the cut along the target path, we can show that the infection rate throughout a segment is smaller than the curing rate. For this reason, during a segment, the process $|D_t|$ has a downward drift. As a consequence, using a standard argument, the expected duration of a segment is small and there is high probability that the segment ends with $|D_t| = 0$, so that the segment is short and we continue with the next segment. As a result, the expected duration of an attempt behaves similar to the case of no infections and is also of order $O(n/r)$. Finally, by studying the number of failed attempts until a successful one, we can establish an upper bound for the overall policy. A formal version of this argument is the content of the rest of this section.

4.3 Segment Analysis

Let us focus on a particular segment, and let $M_t = |D_t|$. The process M_t evolves on the finite set $\{0, 1, \ldots, r/8\Delta\}$. (For simplicity, and without loss of generality, we assume that $r/8\Delta$ is an integer.) Recall that C was defined as the bag on the target path that we were trying to reach at the end of the segment. The difference D_t at the time that the segment starts consists of exactly one node: the target node. Thus, the process M_t is initialized at one, at the beginning of the segment. The process M_t is stopped as soon one of the two boundary points, 0 or $r/8\Delta$, is reached. At each time before the process is stopped, there is a rate equal to r of downward transitions. Furthermore, there is a rate $\mathrm{c}(I_t)$ of upward transitions, corresponding to new infections.

Lemma 2. *The rate* $\mathrm{c}(I_t)$ *of upward transitions during a segment satisfies* $\mathrm{c}(I_t) \leq r/2$.

Proof. The definition $D_t = I_t \setminus C$ implies that $I_t \subseteq C \cup D_t$. Consequently,

$$\begin{aligned} \mathrm{c}(I_t) &\leq \mathrm{c}(C) + \mathrm{c}(D_t) \leq \mathrm{c}(C) + \Delta \cdot |D_t| \\ &= \mathrm{c}(C) + \Delta \cdot M_t \leq \mathrm{c}(C) + \frac{r}{8}. \end{aligned} \tag{4}$$

We have used here Proposition 1, in the first and second inequality, together with the fact $M_t \leq r/8\Delta$.

On the other hand, C is on the target path associated with B, the bag obtained at the end of the waiting period. As remarked at the end of Section 3.1, the impedance does not increase along an optimal crusade, and therefore, $\delta(C) \leq \delta(B)$. Using also Lemma 1, we have

$$\mathrm{c}(C) \leq \delta(C) \leq \delta(B) \leq W + \mathrm{c}(B).$$

Recall now that a waiting period ends with a bag whose cut is at most $r/8$. Therefore, $\mathrm{c}(B) \leq r/8$. It follows that $\mathrm{c}(C) \leq W + r/8$. Using this fact, together with the assumption $r \geq 4W$ and Eq. (4), we obtain

$$\mathrm{c}(I_t) \leq \mathrm{c}(C) + \frac{r}{8} \leq \left(W + \frac{r}{8}\right) + \frac{r}{8} \leq \frac{r}{4} + \frac{r}{8} + \frac{r}{8} = \frac{r}{2}.$$

□

We now establish the properties of the segments that we have claimed earlier; namely, that segments are short, with high probability, and do not last too long.

Lemma 3.

a) *The probability that the segment is long is at most*

$$p = \frac{1}{2^{r/8\Delta} - 1}.$$

b) *The expected length of a segment is upper bounded by* $2/r$.

Proof.

a) Using Lemma 2, the process M_t is stochastically dominated by a process N_t on the same space $\{0, 1, \ldots, r/8\Delta\}$, which is initialized to be equal to the value of M_t at the beginning of the segment (which is one), has a rate r of downward transitions, a rate $r/2$ of upward transitions, and stops at the first time that it reaches one of the two boundary values. Note that the ratio of the downward to the upward drift is equal to 2. The probability, denoted by p, that the process N_t will first reach the upper boundary is a well-studied quantity and is given by the expression in part (a) of the Lemma. The proof is standard and can be found in Section 2.1 of [11] (for a non-martingale based proof) or Section 2.3 of [16] (for a martingale based proof). Since M_t is stochastically dominated by N_t, the probability that M_t will first reach the upper boundary is no larger.

b) For simplicity, let us suppose that the segment starts at time $t = 0$. We define the process

$$H_t = M_t + \frac{r}{2}t$$

and the stopped version, $\hat{H}_t$ which stops at the time T that the segment ends. It is straightforward to verify that $\hat{H}_t$ is a supermartingale, because the upward drift of the process is $\beta\mathrm{c}(I_t) \leq r/2$ and the downward drift is r, so that the total downward drift at least $r/2$. Furthermore, $\hat{H}_0 = H_0 = M_0 = 1$. Using Doob's optional stopping theorem we obtain

$$1 = \mathbb{E}[M_0] = \mathbb{E}[\hat{H}_0] \geq \mathbb{E}[\hat{H}_T] + \frac{r}{2} \cdot \mathbb{E}[T] \geq \frac{r}{2} \cdot \mathbb{E}[T],$$

from which we conclude that

$$\mathbb{E}[T] \leq \frac{2}{r}.$$

□

Note that if $r \geq \alpha \log n$, where α is a sufficiently large constant, then p can be made smaller that $1/n^2$, so that np tends to zero. We will be using this observation later on. We will now bound the length of a waiting period.

Lemma 4. *The expected length of a waiting period is bounded above by $8n/r$.*

Proof. A waiting period involves at most n infections. The waiting period ends as soon as $c(I_t) \leq r/8$. Therefore, during the waiting period, infections happen at a rate of at least $r/8$. In particular, during the waiting period, the expected time between consecutive infections is at most $8/r$. For a maximum of n infections, the expected time is upper bounded by $8n/r$. □

We can now combine the various bounds we have derived so far in order to bound the expected time to extinction under our policy.

Theorem 1. *Suppose that $r \geq 4W$ and that r is large enough so that $np < 1$, where p is as defined in Lemma 3. For any initial bag, the expected time to extinction under the CURE policy is upper bounded by*

$$\frac{1}{1-np} \cdot \frac{10n}{r}.$$

Proof. We start by upper bounding the expected duration of an attempt. The expected length of the waiting period of an attempt is upper bounded by $8n/r$, by Lemma 4.

The number of segments during an attempt is at most n since each segment is associated with one target node and there can be at most n different target nodes. By Lemma 3, the expected length of a segment is at most $2/r$.

Putting everything together, the expected duration of an attempt is at most $(8n/r) + (2n/r) = 10n/r$.

Each attempt involves n segments. During each segment, there is probability at most p that the segment is long and that the attempt fails. Therefore, the overall probability that an attempt will fail is at most np (here we used the union bound). We note that his upper bound (np) on the failure probability holds regardless of the initial bag at the beginning of an attempt. It follows that the attempt is stochastically dominated by a geometric random variable with parameter $1 - np$. For this reason, the expected number of attempts is at most $1/(1 - np)$, and the desired result follows. □

5 Corollaries and Near-Optimality of the CURE Policy

Theorem 1 has a number of interesting consequences, which we collect in the corollary that follows. We argue that if all nodes are initially infected, then the expected time to extinction under any policy is at least n/r. Furthermore, in a certain regime of parameters, our policy achieves $O(n/r)$ expected time to extinction and is therefore optimal within a multiplicative constant. Finally, if the CutWidth increases sublinearly with the number of nodes, then the expected time to extinction can be made sublinear in n, using only a sublinear budget. This last result is also proved in [6], using a different, nonconstructive argument.

Corollary 1.

a) *For any graph with n nodes and with all nodes initially infected, the expected time to extinction is at least n/r, under any policy.*

b) *Suppose that the budget r satisfies*

$$r \geq 4W, \qquad r \geq 16\,\Delta \log_2 n.$$

Then, for large enough n, and for any initial set of infected nodes, the expected time to extinction under the CURE policy is at most $26n/r$, which is sublinear in n and within a multiplicative factor from optimal.

c) *Suppose that the maximum degree is bounded, i.e., Δ is $O(1)$. If the CutWidth increases sublinearly with n, then it is possible to have sublinear time to extinction with a sublinear budget.*

Proof.

a) Since nodes are cured at a rate of at most r, and there are n nodes to be cured, the expected time to extinction must be at least n/r, even in the absence of infections.

b) When $r \geq 16 \cdot \log_2 n \cdot \Delta$, we have $r/8\Delta \geq 2 \log_2 n$, and $2^{r/8\Delta} \geq n^2$. Thus, the probability p in Lemma 3 is of order $O(1/n^2)$, and np is of order $O(1/n)$. In particular, for large enough n, the factor $1/(1 - np)$ is less than 2. By Theorem 1, the expected time to extinction is at most $20n/r$. This is sublinear in n, because r tends to infinity. Order optimality follows from part (a).

c) Suppose that the budget r satisfies the conditions in part (b), together with the condition

$$r = \Omega(n/\log n).$$

Then, it follows from part (b) that the expected time to extinction under the CURE policy is of order $O(\log n)$. If W increases sublinearly with n, we can satisfy the conditions in parts (b) and (c) while keeping r sublinear in n, and still achieve sublinear, e.g., $O(\log n)$ expected time to extinction. □

We continue with some examples. For a line graph with n nodes, the CutWidth is equal to 1 and $\Delta = 2$. Therefore, by part (b) of Corollary 1 we can guarantee an approximately optimal expected time to extinction, of order $O(n/r)$, as long as $r \geq 16 \cdot \log_2 n \cdot \Delta = 32 \log_2 n$. We note, however, that for this example, our analysis is not tight, and the requirement $r \geq 32 \log_2 n$ is stronger than necessary.

For a square grid-graph with n nodes, the Cut-Width is approximately $\sqrt{n}$ and $\Delta = 4$. In this case, the requirement $r \geq 4W \approx 4\sqrt{n}$ is the dominant one, and suffices to guarantee an approximately optimal expected time to extinction, of order $O(n/r)$.

In both of these examples, we can of course let r be much larger than the minimum required, which was $O(\log n)$ and $O(\sqrt{n})$, respectively, in order to obtain a smaller expected time to extinction, e.g., the $O(\log n)$ expected time to extinction in part (c) of the corollary.

6 Performance of the CURE Policy under Arbitrary Initial Infections

The results of Section 5 are stated in terms of n and W which are global characteristics of the network and do not take into account the possibility of a favorable set of initially infected nodes. In this section we obtain performance guarantees for our policy as a function of $|A|$ and $\delta(A)$, where A is the bag of initially infected nodes. Our goal is to explore conditions under which the CURE policy is (order) optimal, i.e., achieves expected extinction time of order $O(|A|/r)$.

Note that if $\mathrm{c}(A) > r/8$, a waiting phase is initiated. By the end of the waiting phase a superset of A (potentially the whole graph) is infected and thus the performance of the CURE policy cannot be related to the properties of A. For this reason, we focus on the case where $\mathrm{c}(A) < r/8$. Section 5 illustrates that when the budget is larger than $4W$ then, the CuRe policy is (order) optimal. In this section we are interested in the case where the impedance of the initial bag, $\delta(A)$, is smaller than the CutWidth of the graph, i.e., $\delta(A) < W$. Under such conditions, we expect to require less curing budget in order to attain (order) optimal extinction time; the main theorem of this section confirms this fact.

First we establish some properties of the *first* attempt of the CURE policy, when $r \geq \max\{4\delta(A), 8\mathrm{c}(A)\}$. Note the similarity between the latter condition and that of Corollary 1(a).

Lemma 5. *Suppose that the set of initially infected nodes is A, and that $r \geq \max\{4\delta(A), 8\mathrm{c}(A)\}$. Let τ_S denote the duration of a segment and let $\mathcal{S}$ denote the event that the segment is short. Moreover, we write $p_l = \mathbb{P}(\mathcal{S}^c)$. Then, for the first attempt the following properties hold:*

a) *The probability p_l that a segment is long is at most*

$$p = \frac{1}{2^{r/8\Delta} - 1}.$$

b) *The expected length of a segment is upper bounded by $2/r$, i.e., $\mathbb{E}[\tau_s] \leq 2/r$.*

c) *The conditional expectation of a segment, given that it is short, $\mathbb{E}[\tau_s \,|\, S]$, is upper bounded by $2/(r(1-p))$.*

Proof.

a,b) Note that since $\mathrm{c}(A) \leq r/8$ there is no waiting phase and the target path of the first attempt is the crusade associated with $\delta(A)$. Given this observation, the proofs are identical to Lemma 3 after replacing W by $\delta(A)$ in all arguments.

c) We have,

$$\begin{aligned}\mathbb{E}[\tau_s] &= \mathbb{E}[\tau_s \,|\, \mathcal{S}](1-p_l) + \mathbb{E}[\tau_s \,|\, \mathcal{S}^c]p_l \\ &\geq \mathbb{E}[\tau_s \,|\, \mathcal{S}](1-p_l) \geq \mathbb{E}[\tau_s \,|\, \mathcal{S}](1-p).\end{aligned}$$

Solving for $\mathbb{E}[\tau_s \,|\, \mathcal{S}]$ and using part (b) the result follows. □

We now combine the bounds we derived in order to bound the expected time to extinction under our policy.

Lemma 6. *Suppose that the set of initially infected nodes is A with $r \geq \max\{4\delta(A), 8\mathrm{c}(A)\}$. Moreover, suppose that r is large enough so that $|A|p < 1$ and let $\mathcal{E}$ denote the event that the first attempt is successful. Then*

$$\mathbb{E}[\tau \,|\, \mathcal{E}] \leq |A| \frac{2}{(1-p)r}.$$

Proof. First, the conditional expectation is well defined since $\mathbb{P}(\mathcal{E}) \geq 1 - |A|p > 0$ by the assumptions of the lemma. Conditioned on the success of the first attempt, the number of segments is $|A|$ and the result follows from Lemma 5c. □

Lemma 6 is mainly relevant in the regime where $|A|$ grows to infinity with

$$r \geq \max\{4\delta(A), 16\,\Delta \log_2|A|, 8\mathrm{c}(A)\}. \tag{5}$$

In this regime, the budget is sufficiently high for the first attempt to be successful with high probability. Thus, the performance indicated by Lemma 6 is achieved conditioned on an event which occurs with high probability, as the following theorem states.

Theorem 2. *Suppose that the budget satisfies Eq. (5) and that the set of initially infected nodes is A, whose size $|A|$ grows to infinity. Let $\mathcal{E}$ be the event that the first attempt is successful. Then, $\mathbb{P}(\mathcal{E}) = 1 - o(1)$, $\mathbb{E}[\tau \,|\, \mathcal{E}]$ is of order $O(|A|/r)$, and thus our policy is (order) optimal with high probability.*

Proof. Following similar reasoning as in Corollary 1, under the condition (5), the probability p in Lemma 5 is of order $O(1/|A|^2)$. This implies that

$$\lim_{|A|\to\infty} \mathbb{P}(\mathcal{E}) \geq \lim_{|A|\to\infty} (1 - |A|p) = 1.$$

Moreover, for large enough $|A|$, $1-p$ is larger than $1/2$ and thus, by Lemma 6 the expected time to extinction, conditioned on $\mathcal{E}$ is at most $4|A|/r$ and thus $O(|A|/r)$. □

Note that Theorem 2 establishes (order) optimality with high probability, which is weaker than (order) optimality in Corollary 1. This is due to the fact that the lower budget requirements $(r \geq \max\{4\delta(A), 16\,\Delta \log_2|A|, 8\mathrm{c}(A)\}$ versus $r \geq \max\{4W, 16 \log_2 n \cdot \Delta\})$ come at a cost: if we have a long segment and a failed attempt (which is a small probability event) the process can potentially be uncontrollable and the extinction time from then on large.

7 Discussion and Conclusions

We have presented a dynamic curing policy which achieves sublinear expected time to extinction, using a sublinear curing budget when the CutWidth of the underlying graph is sublinear in the number of nodes. This policy applies to any subset of initially infected nodes and the resulting expected time to extinction is order-optimal when the available budget is sufficiently large.

The analysis of the extinction time under our policy is based on a drift analysis of the epidemic process. The upward drift is equal to the cut of the set of infected nodes $\mathrm{c}(I_t)$ and the downward drift is proportional to the curing budget r. While the process is on the target path, $\mathrm{c}(I_t)$, and therefore the upward drift, can be bounded from above by the impedance of the starting bag. On the other hand, when the process deviates from the target path this is no longer the case. For this reason we invoke the maximum degree Δ of the graph in order to bound the change of the cut during each such deviation. Note that none of our results (except for Corollary 1c) requires bounded degree. The maximum degree appears in the minimum budget requirement but is not required to be bounded. Furthermore, our results indicate that under our policy, the process has low probability of deviating significantly from the target path and therefore only the locally maximum degree is relevant to the analysis, and not the global maximum. In other words, as long as the infection does not reach high degree nodes we should have results similar to those for the bounded degree case. However, the performance analysis for this case is expected to be significantly harder and the statement of the results more complicated.

Our policy allocates all the available budget to one node at every time instant. This is permitted by our formulation but in practice each infected agent can only be offered a bounded amount of curing resources. Our policy, cannot be directly generalized to account for such a constraint but the insights of our solution can be directly adapted to such a scenario.

A drawback of the CURE policy is computational complexity because calculating the impedance of a bag or finding a target path is computationally hard. Like many other interesting graph problems, CutWidth is NP-complete [7], even if we restrict to planar graphs or graphs with maximum degree three [13] but in general fixed parameter linear [17]. Several approximation algorithms have been developed for computing the CutWidth of a graph. Specifically, there is a polynomial time $O(\log^2 n)$-approximation algorithm for general graphs [10], and a polynomial time constant factor approximation algorithm for dense graphs [15]. We leave it as an interesting future direction to develop such algorithms for computing the impedance of a bag. Finally, we have argued in this paper that the CURE policy is efficient in the sense of attaining near-optimal, $O(n/r)$ expected time to extinction, in a certain parameter regime. It is an interesting problem to look for approximately optimal policies over a wider set of regimes.

ACKNOWLEDGMENTS

Research partially supported by the Draper Laboratories and NSF grant CMMI-1234062. A preliminary version of some of the results in this paper are included in a paper submitted to the 53rd IEEE Conference on Decision and Control, December 2014. K. Drakopoulos is the corresponding author.

REFERENCES

[1] D. Bienstock and P. Seymour. (1991). Monotonicity in graph searching. *J. Algorithms* [Online]. *12(2)*, pp. 239–245. Available: http://www.sciencedirect.com/science/article/pii/019667749190003H

[2] C. Borgs, J. Chayes, A. Ganesh, and A. Saberi. (2010). How to distribute antidote to control epidemics. *Random Structures and Algorithms* [Online]. *37(2)*, pp. 204–222. Available: http://dx.doi.org/10.1002/rsa.20315

[3] F. R. K. Chung, P. Horn, and A. Tsiatas. (2009). Distributing antidote using pagerank vectors. *Internet Math.* [Online]. *6(2)*, pp. 237–254. Available: http://dblp.uni-trier.de/db/journals/im/im6.html#ChungHT09

[4] M. Chung, F. Makedon, I. Sudborough, and J. Turner. (1985). Polynomial time algorithms for the min cut problem on degree restricted trees. *SIAM J. Comput.* [Online]. *14(1)*, pp. 158–177. Available: http://epubs.siam.org/doi/abs/10.1137/0214013

[5] R. Cohen, S. Havlin, and D. ben-Avraham, "Efficient immunization strategies for computer networks and populations," *Phys. Rev. Lett.*, vol. 91, p. 247901, 2003.

[6] K. Drakopoulos, A. Ozdaglar, and J. N. Tsitsiklis, *Working Paper*, 2015.

[7] F. Gavril, "Some NP-complete problems on graphs," *11th Conf. Inf. Sci. Syst.*, 1977, pp. 91–95.

[8] E. Gourdin, J. Omic, and P. Van Mieghem. (2011, Oct.). Optimization of network protection against virus spread. in *Proc. 8th Int. Workshop Des. Rel. Commun. Netw.*, pp. 86–93. [Online]. Available: http://dx.doi.org/10.1109/drcn.2011.6076889

[9] E. Korach and N. Solel. (1993). Tree-width, path-width, and cutwidth. *Discrete Appl. Math.* [Online]. *43(1)*, pp. 97–101. Available: http://www.sciencedirect.com/science/article/pii/0166218X9390171J

[10] T. Leighton and R. Satish. (1999, Nov.). Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM* [Online]. *46(6)*, pp. 787–832. Available: http://doi.acm.org/10.1145/331524.331526

[11] D. A. Levin, Y. Peres, and E. L. Wilmer. (2006). *Markov Chains and Mixing Times* [Online]. American Mathematical Society. Available: http://scholar.google.com/scholar.bib?q=info:3wf9IU94tyMJ:scholar.google.com/&output=citation&hl=en&as_sdt=2000&ct=citation&cd=0

[12] F. Makedon and I. Sudborough. (1989). On minimizing width in linear layouts. *Discrete Applied Mathematics* [Online]. *23(3)*, pp. 243–265. Available: http://www.sciencedirect.com/science/article/pii/0166218X89900164

[13] B. Monien and I. Sudborough. (1988). Min cut is NP-complete for edge weighted trees. *Theoretical Comput. Sci.* [Online]. *58(1–3)*, pp. 209–229. Available: http://www.sciencedirect.com/science/article/pii/030439758890028X

[14] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas. (2013). Optimal vaccine allocation to control epidemic outbreaks in arbitrary networks. *CoRR*, vol. abs/1303.3984 [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr1303.html#abs-1303-3984

[15] A. Sanjeev, F. Alan, and K. Haim. (2002). A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. *Mathematical Programming* [Online]. *92(1)*, pp. 1–36. Available: http://dx.doi.org/10.1007/s101070100271

[16] J. M. Steele, *Stochastic Calculus and Financial Applications*. New York, NY, USA: Springer, 2000.

[17] D. M. Thilikos, M. Serna, and H. L. Bodlaender. (2005, Jul.). Cutwidth I: A Linear Time Fixed Parameter Algorithm. *J. Algorithms* [Online]. *56(1)*, pp. 1–24. Available: http://dx.doi.org/10.1016/j.jalgor.2004.12.001

Kimon Drakopoulos received the diploma in electrical and computer engineering from the National Technical University of Athens, Athens, Greece, in 2009 and the MSc degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, in 2011. From 2011 to present, he is working towards the PhD degree at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. His current research interests include social network analysis, network science, applied probability, game theory, and network economics. He is a member of the IEEE.

Asu Ozdaglar received the BS degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 1996, and the SM and the PhD degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1998 and 2003, respectively. She is currently a professor in the Electrical Engineering and Computer Science Department at the Massachusetts Institute of Technology and the director of the Laboratory for Information and Decision Systems. She is also a member of the Operations Research Center. Her research expertise includes optimization theory, with emphasis on nonlinear programming and convex analysis, game theory, with applications in communication, social, and economic networks, distributed optimization and control, and network analysis with special emphasis on contagious processes, systemic risk and dynamic control. She is the recipient of a Microsoft fellowship, the MIT Graduate Student Council Teaching award, the NSF Career award, the 2008 Donald P. Eckman award of the American Automatic Control Council, the Class of 1943 Career Development Chair, the inaugural Steven and Renee Finn Innovation Fellowship, and the 2014 Spira teaching award. She served as a member of the Board of Governors of the Control System Society and as an associate editor for the *IEEE Transactions on Automatic Control*. She is currently the area co-editor for a new area for the journal Operations Research, entitled "*Games, Information and Networks*" and the chair of the Control System Society Technical Committee "*Networks and Communication Systems*". She is the co-author of the book entitled "*Convex Analysis and Optimization*" (Athena Scientific, 2003).

John N. Tsitsiklis received the BS degree in mathematics and the BS, MS, and PhD degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1980, 1980, 1981, and 1984, respectively. He is currently a Clarence J. Lebel professor with the Department of Electrical Engineering and Computer Science at MIT. He has served as a co-director of the MIT Operations Research Center from 2002 to 2005, in the National Council on Research and Technology in Greece (2005-2007), and currently chairs the Council of the Harokopio University in Athens, Greece. His research interests are in systems, optimization, communications, control, and operations research. He has coauthored four books and several journal papers in these areas. He has been a recipient of an Outstanding Paper Award from the IEEE Control Systems Society (1986), the M.I.T. Edgerton Faculty Achievement Award (1989), the Bodossakis Foundation Prize (1995), an ACM Sigmetrics Best Paper Award (2013), and a co-recipient of two INFORMS Computing Society prizes (1997, 2012). He is a member of the National Academy of Engineering. In 2008, he was conferred the title of Doctor honoris causa, from the Universit Catholique de Louvain. He is a fellow of the IEEE.