# Supplement to “Sampling Strategies for Epidemic-Style Information Dissemination”[1]

Varun Gupta and Milan Vojnović

June 2009

Technical Report
MSR-TR-2009-70

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
`http://www.research.microsoft.com`

[1]This is a supplement to M. Vojnović, V. Gupta, T. Karagiannis, C. Gkantsidis, “Strategies for Epidemic-Style Information Dissemination”, in Proc. of IEEE Infocom 2008, Phoenix, AZ, USA. Varun Gupta is a doctoral student in the Computer Science Department, Carnegie Mellon University; Milan Vojnović is a researcher with the Systems and Networking group at Microsoft Research, Cambridge (milanv@microsoft.com).

**Abstract – In this note we establish several new results for sampling-based information dissemination strategies considered in [1] (version with proofs – [2]). First, we establish the computational complexity and an algorithm for computing the target set of subnets for the optimum static sampling strategy, OPT-STATIC [1], given the densities of susceptible hosts over subnets. Second, we identify the worst-case and best-case placements of initially infected hosts, so as to maximize or minimize, respectively, the number of samplings required to reach a given fraction of susceptible hosts. Finally, we study the effect of target fraction of infected hosts on the optimum number of samplings, and consider simple approximations for the optimum number of samplings to reach a given target fraction of hosts. The surprising accuracy of our simple approximation provides insights on the qualitative and quantitative impact of system parameters on the optimum number of samplings. We illustrate some of our results using the empirical distributions of hosts over subnets that were used in [1].**

# 1 Introduction

In this note we provide several new results on sampling-based information dissemination strategies considered in [1] (version with proofs [2]). We summarize the setting and some of the results from [1] which we use in this note.

## 1.1 Model and Notation

We consider an address space of size $\Omega$ partitioned into $J$ subnets. Subnet $j$ occupies a fraction $\omega_j \geq 0$ of the address space. Of the $\Omega$ addresses, $N$ addresses are occupied by vulnerable (or susceptible) hosts. The fraction of vulnerable hosts residing in subnet $j$ is denoted by $n_j$. We use $I_j(t)$ and $S_j(t)$ to denote the number of infected and susceptible hosts, respectively, in subnet $j$ at time $t \geq 0$. We use $i_j(t) = \frac{I_j(t)}{N}$ and $s_j(t) = \frac{S_j(t)}{N}$ to denote the normalized versions of the corresponding variables. We denote with $i(0)$ the initial fraction of infected hosts over all subnets, thus $i(0) = \sum_{j=1}^{J} i_j(0)$. We use the following vector notation $\vec{n} = (n_1, \dots, n_J)$, $\vec{s}(t) = (s_1(t), \dots, s_J(t))$, and $\vec{\omega} = (\omega_1, \dots, \omega_J)$. We assume that $\vec{n}$ and $\vec{s}(0)$ may be known, but the identities (addresses) of these hosts is unknown.

For each subnet $j$, let $d_j = s_j(0)/\omega_j$, and let $((1), (2), \dots, (J))$ denote a sorted permutation of the subnet indices so that the following holds:

$$d_{(1)} \geq d_{(2)} \geq \cdots \geq d_{(J)}.$$

## 1.2 Summary of Optimal Sampling Strategy

In [1, Theorem 2, Corollary 6], it was established that for a given initial fraction of infected hosts, $i(0)$, the minimum number of samplings per host, $u$ ($= \int_0^\infty i(t)\mathbf{1}_{i(t)<i^0} dt$), to reach a given final fraction of infected hosts, $i^0$, is given by:

$$u(\vec{s}(0), \vec{\omega}, i^0) = \frac{\sum_{k=1}^{J'} \omega_{(k)}}{\beta} \left[ \log\left( \frac{1}{1 - \frac{i^0 - i(0)}{\sum_{k=1}^{J'} s_{(k)}(0)}} \right) - D(\vec{\omega}^{J'} || \vec{s}^{J'}(0)) \right] \quad (1)$$

where $\beta = \frac{N}{\Omega}$ denotes the number of vulnerable hosts per address,

$$J' = \max\left\{ j : \frac{s_{(j)}(0)}{\omega_{(j)}} > \frac{\sum_{k=1}^{j} s_{(k)}(0) - (i^0 - i(0))}{\sum_{k=1}^{j} \omega_{(k)}} \right\}, \tag{2}$$

and $D(\vec{\omega}^{J'} || \vec{s}^{J'}(0))$ is the Kullback-Liebler divergence of the following two distributions

$$\vec{\omega}^{J'} = \left( \frac{\omega_{(j)}}{\sum_{k=1}^{J'} \omega_{(k)}},\ j = 1, \ldots, J' \right),\ \vec{s}^{J'} = \left( \frac{s_{(j)}(0)}{\sum_{k=1}^{J'} s_{(k)}(0)},\ j = 1, \ldots, J' \right).$$

From (1), we have that $f(\vec{d}) := \beta u$ can be written in the following form

$$f(\vec{d}) = \sum_{k=1}^{J'} \omega_{(k)} \log d_{(k)} - \left( \sum_{k=1}^{J'} \omega_{(k)} \right) \log \left( B_{J'}(\vec{d}) \right) \tag{3}$$

where,

$$B_j(\vec{d}) = \frac{s^0 - \sum_{k=j+1}^{J} \omega_{(k)} d_{(k)}}{\sum_{k=1}^{j} \omega_{(k)}}$$

denotes the final density of vulnerable hosts within the target set, if the target set is $\{(1), \ldots, (j)\}$ and the sampling probabilities are chosen optimally. Using the above notation, we can rewrite (2) as

$$J' = \max\left\{ j : d_{(j)} > B_j(\vec{d}) \right\}. \tag{4}$$

Let $A = \{(1), \ldots, (J')\}$. Intuitively, the above equation characterizes the optimal target set $A$ as the largest set of subnets such that the initial density of vulnerable hosts in each of the subnets in $A$ is larger than the final density of vulnerable hosts in the set $A$.

It is further shown [1, Corollary 5] that the optimum number of samplings per host (1) can be achieved by sampling over the subnets of the target set $A$ with either time-invariant (OPT-STATIC), or time-varying probabilities of sampling a subnet in the target set $A$. While there are multiple dynamic strategies which are optimal with respect to the number of samplings, the following intuitive dynamic strategy is also optimal with respect to the time to reach the target population:

OPT-DYNAMIC
At any time $t \geq 0$, each infected host samples an address uniformly at random from the address space of the subnets in the set $S(t)$, where $S(t)$ contains the densest subnets with respect to the densities $s_j(t)/\omega_j$.

We will use the equivalence between the number of samplings for the above dynamic strategy and the number of samplings for the optimal static scheme to establish results regarding the latter.

## 1.3 Structure of this Note

In Section 2 we establish the computational complexity and an algorithm for computing the optimum set of subnets to target with scans, given the densities of susceptible hosts per subnet. Section 3 establishes various properties for the optimum number of samplings.

## 2 Computation of the Target Set

We begin by considering a generalization of the problem of computing the optimal target set $A$: Consider a set of items $I$ such that each item $i_j \in I$ is is associated with a value $v_j$, and weight $w_j$. Let $x$ be an input parameter, $x \in [0, 1]$. The problem TOP-SET is defined as follows:

TOP-SET
Find $S \subseteq I$ such that $i_j \in S$ if and only if

$$v_j > \frac{\sum_{k \in S} w_k v_k - x}{\sum_{k \in S} w_k}.$$

The computational complexity of solving TOP-SET is specified by the following result.

**Proposition 1.** *The time-complexity of problem* TOP-SET *is* $\Theta(|I|)$. *Further, problem* TOP-SET *is solved by algorithm* FIND TOP-SET *in* $O(|I|)$ *time.*

*Proof.* To solve TOP-SET one must inspect each item in the set $I$, hence we need $\Omega(|I|)$ time. The result follows by the algorithm below that solves the problem in $O(|I|)$ time. □

**Algorithm** FIND TOP-SET
**Input**: Set of items $I = \{i_1, \ldots, i_J\}$
$v_j$ = value of item $i_j$
$w_j$ = weight of item $i_j$
$x$ = target volume
**Output**: $S \subseteq I$ such that $i_j \in S \iff v_j \geq \frac{\sum_{k \in S} w_k \cdot v_k - x}{\sum_{k \in S} w_k}$
**Initialize**: $S = \emptyset$, $T = I$, $a = b = 0$
**while** $T \neq \emptyset$ **do**
  $m \leftarrow \text{median}\left(\{v_j | i_j \in T\}\right)$ ;
  $L \leftarrow \{i_j | i_j \in T, v_j > m\}$ ;
  $R \leftarrow \{i_j | i_j \in T, v_j < m\}$ ;
  $M \leftarrow T \setminus (L \cup R)$ ;
  $a' \leftarrow \sum_{k \in L \cup M} w_k \cdot v_k$ ;
  $b' \leftarrow \sum_{k \in L \cup M} w_k$ ;
  **if** $m \geq \frac{a + a' - x}{b + b'}$ **then**
    $a \leftarrow a + a'$ ;
    $b \leftarrow b + b'$;
    $S \leftarrow S \cup L \cup M$ ;
    $T \leftarrow R$
  **else**
    $T \leftarrow L$ ;
  **end**
**end**

Each iteration of the while loop in the above algorithm takes $O(|T|)$ time (there is a deterministic linear time algorithm for finding median, as well as a much simpler randomized algorithm that runs in linear time with high probability [3, Chapter 9]), and we reduce the size of $|T|$ by at least half in each iteration. Therefore, algorithm FIND TOP-SET runs in $O(|I|)$ time.

**Finding the optimum target set $A$.** Define the item values as $v_j = d_j$, weights as $w_j = \omega_j$, and let the input parameter be $x = (i^0 - i(0))$.

# 3 Properties of the Optimum Number of Samplings

In this section we examine the dependence of the optimum number of samplings on the system parameters via bounds and approximations. In Section 3.1, we present results regarding the placement of initially infected nodes that maximize or minimize the optimum number of samplings, and thus give us bounds on the effect of placement of initially infected nodes. In Section 3.2, we present results characterizing the effect of target fraction of infected nodes on the optimum number of samplings. In Section 3.3, we consider the special case where the initially infected hosts are uniformly placed among the vulnerable hosts. Finally, Section 3.4 presents simple approximations and bounds for the optimum number of samplings.

## 3.1 Placement of Initially Infected Hosts

In this section, we identify the worst-case and best-case placement of initially infected hosts that, respectively, maximize and minimize the optimum number of samplings.

### 3.1.1 Worst Placement

We consider the maximum optimum number of samplings over all distributions of infected hosts over subnets given that the total fraction of infected hosts is $i(0)$. We want to solve the following problem

$$\begin{array}{ll} \text{MAX-SAM}: & \\ \text{maximize} & f(\vec{d}) \\ \text{over} & d_{(k)} \in \left[0, \frac{n_{(k)}}{\omega_{(k)}}\right],\ k = 1, \ldots, J \\ \text{subject to} & \sum_{k=1}^{J} \omega_{(k)} d_{(k)} = s(0) \end{array}$$

where $f(\vec{d})$ is given by (3).

**Proposition 2.** *The initial placement of the infected hosts over subnets that maximizes the number of samplings is obtained by starting with the system with no infected hosts and then continually adding infected hosts to most dense subnets (where density of a subnet $j$ is $d_{(j)}$).*

In what follows we provide a simple intuitive proof of the above proposition. We provide an alternate proof in Appendix A, which establishes additional structural properties regarding the OPT-STATIC strategy.

*Proof.* To prove the proposition we will make use of [1, Corollary 5] that implies that for any given target fraction of infected hosts $i^0 \in [i(0), 1]$ the schemes OPT-DYNAMIC and OPT-STATIC require the same number of samplings which is smallest possible. Recall that OPT-DYNAMIC corresponds to sampling the currently densest subnets at all times, where density of a subnet $j$ at time $t \geq 0$ is equal to $s_{(j)}(t)/\omega_j$.

For OPT-DYNAMIC scheme, the number of samplings to reach the target fraction of infected hosts $i^0$ can be represented as

$$u(i^0) = \frac{1}{\beta} \int_{i(0)}^{i^0} \frac{di}{\delta(i)}$$

where $\delta(i)$ is the density of a densest subnet when the total fraction of infected hosts is $i$. Indeed, this follows from

$$\frac{d}{dt} i(t) = \beta \delta(i(t)) i(t)$$

and then noting that

$$u(t) = \int_0^t i(v) dv = \int_0^t \frac{1}{\beta} \frac{di(v)}{\delta(i(v))}$$

from which the assertion follows.

Now, for any given $i^0$, $u(i^0)$ is maximized if $\delta(i)$ is minimum for each $i \in [i(0), i^0]$. This is clearly achieved by starting with no infected hosts, and simulating the OPT-DYNAMIC strategy for first $i(0)$ fraction of infected hosts. That is, by adding the initially infected hosts to the densest subnets with respect to $n_{(j)}/\omega_{(j)}$, $j = 1, \dots, J$. The result then follows from [1, Corollary 5], as the minimum $u(i^0)$ under OPT-DYNAMIC can be achieved under OPT-STATIC.

□

#### 3.1.2 Best Placement

We consider the minimum optimum number of samplings over all distributions of initially infected hosts over subnets given that the total fraction of initially infected hosts is $i(0)$. Formally, we consider

$$\begin{array}{ll} \text{MIN-SAM}: & \\ \text{minimize} & f(\vec{d}) \\ \text{over} & d_{(k)} \in \left[0, \frac{n_{(k)}}{\omega_{(k)}}\right], \ k = 1, \dots, J \\ \text{subject to} & \sum_{k=1}^{J} \omega_{(k)} d_{(k)} = s(0). \end{array}$$

**Proposition 3.** *The initial placement of the infected hosts over subnets that minimizes the number of samplings is obtained by starting with the system with no infected hosts and then continually adding infected hosts to least dense subnets (where density of a subnet $j$ is $d_{(j)}$).*

*Proof.* We proceed exactly as in proof of Proposition 2, using the observation that for any given target fraction of infected hosts $i^0 \in [i(0), 1]$ the schemes OPT-DYNAMIC and OPT-STATIC require the same number of samplings. Further, for any given $i^0$, the optimum number of samplings under OPT-DYNAMIC is minimized if $\delta(i)$, the density of the densest subnets when the fraction of infected hosts is $i$, is maximum for each $i \in [i(0), i^0]$. This is clearly achieved by assigning all the initially infected hosts to sparsest subnets with respect to $n_{(j)}/\omega_{(j)}$, $j = 1, \dots, J$. □

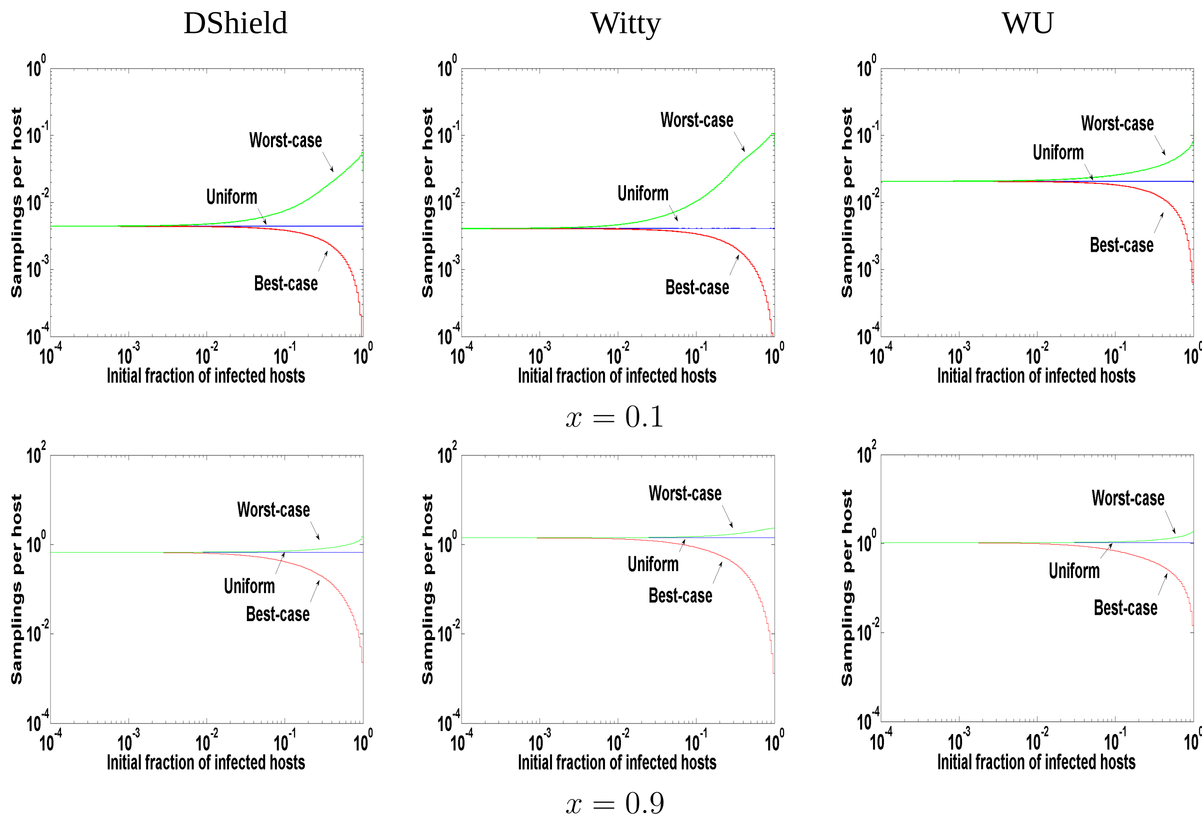



Figure 1: The number of samplings per host vs. the initial fraction of hosts for worst-case, best-case, and uniform placement (see Section 3.3) of initially infected hosts. The target fraction of infected hosts is set to $i(0) + x(1 - i(0))$, where $x = 0.1$ (top row) and $x = 0.9$ (bottom row).

### 3.1.3 Numerical Results

In Fig. 1 we show the results for the optimum number of samplings versus the initial fraction of infected hosts for three different placements of infected hosts over subnets. The results suggest that for realistic distributions of vulnerable hosts over subnets, the placement of initially infected hosts may have significant impact only if the total fraction of initially infected hosts is sufficiently large (e.g., order 1 out of 100 hosts, or more). Further, for large target population ($x = 0.9$), while the performance of best-case placement approaches that of uniform placement for a large range of values of the fraction of initially infected hosts, best-case placement still provides appreciable gains for large $i(0)$.

## 3.2 Effect of Final Fraction of Infected Hosts

**Proposition 4.** *The optimum number of samplings per host, $u$ (given in Eq. (1)), is an increasing, convex function of the final fraction of infected hosts $i^0$.*

*Proof.* We again use the equivalence between the optimum number of samplings for strategies

OPT-STATIC and OPT-DYNAMIC. From the proof of Proposition 2, we have:

$$\frac{\partial u}{\partial i^0} = \frac{1}{\beta\delta(i^0)}$$

where, recall, $\delta(i^0)$ denotes the density of susceptible hosts in the densest subnets under OPT-DYNAMIC when the fraction of infected population is $i^0$. It is not hard to see that $\delta(i^0)$ is a positive, continuous, decreasing, piece-wise linear, convex function. Therefore, $(\partial/\partial i^0)u$ is a positive, increasing function and the proposition follows.

□

### 3.3 Uniform Distribution of Initially Infected Hosts

In this section, we consider the special case where initially infected hosts are placed uniformly among the vulnerable hosts. This is a natural assumption – it holds when, initially, each vulnerable host is equally likely to be infected. In this case, we have

$$s_j(0) = (1 - i(0))n_j,\ j = 1, \ldots, J.$$

Under the above assumption, we have the following result characterizes the optimum number of samplings.

**Proposition 5.** *Under assumption (A), the optimum number of samplings $u$ given by Eq. (1) is a function of three parameters $(x, \vec{n}, \vec{\omega})$ where $x$ is the relative decrease of the number of susceptibles, i.e. $x = (i^0 - i(0))/(1 - i(0))$. Specifically, we have*

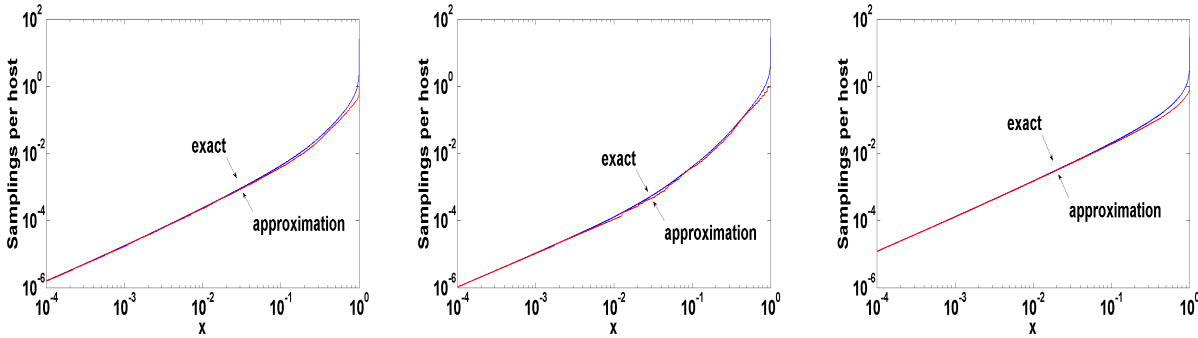
$$\begin{aligned} u(x, \vec{n}, \vec{\omega}) &= \frac{\sum_{k=1}^{J'} \omega_{(k)}}{\beta}\left[\log\left(\frac{1}{1 - \frac{x}{\sum_{k=1}^{J'} n_{(k)}}}\right) - D(\vec{\omega}^{J'} || \vec{n}^{J'})\right] \\ J' &= \max\left\{j : \frac{n_{(j)}}{\omega_{(j)}} > \frac{\sum_{k=1}^{j} n_{(k)} - x}{\sum_{k=1}^{j} \omega_{(k)}}\right\} \end{aligned}$$

*where* $\vec{n}^{J'} = \left(\frac{n_{(j)}}{\sum_{k=1}^{J'} n_{(k)}},\ j = 1, \ldots, J'\right)$.

Note that Proposition 5 says that if, initially, the infected hosts are placed uniformly, then the initial fraction of infected hosts, $i(0)$, and the final fraction of infected hosts, $i^0$, influence the optimum number of samplings only via the ratio $x = \frac{i^0 - i(0)}{1 - i(0)}$. This is also highlighted in Figure 1, where the optimum number of samplings is independent of $i(0)$ under uniform placement.

From Proposition 4, we know that for a fixed $i(0)$, the optimum number of samplings is an increasing convex function of $i^0$, and hence $x$. However, using Proposition 5, we further have the following.

**Corollary 1.** *The optimum number of samplings per subnet $u(x, \vec{n}, \vec{\omega})$ is an increasing, convex function of $x$.*




(dshield) (witty) (wu)

Figure 2: The number of samplings per host vs. $x$ where $x = (i^0 - i(0))/(1 - i(0))$.

### 3.4 Lower Bounds and Approximations

**Proposition 6.** *The total number of samplings per host to infect the fraction of hosts $i^0$ satisfies*

$$u \geq \frac{1}{\beta \rho_{J'}}(i^0 - i(0)) - D(\vec{\omega}^{J'}||\vec{s}^{J'}(0))$$

*where $\rho_{J'}$ is the initial density of susceptible hosts over the target set, i.e.*

$$\rho_{J'} = \frac{\sum_{k=1}^{J'} s_{(k)}(0)}{\sum_{k=1}^{J'} \omega_{(k)}}.$$

*Proof.* The result follows directly from (1) and using the inequality $-\log(1-x) \geq x$ for $x < 1$. $\square$

In [1], we presented empirical evidence for the fact that under real world distributions for hosts, the KL-divergence term ($D(\vec{\omega}^{J'}||\vec{s}^{J'})$) is negligible compared to the leading term in the expression for optimum number of samplings, $u$. We thus propose the following approximation:

$$u \approx \frac{1}{\beta \rho_{J'}}(i^0 - i(0)). \tag{5}$$

Note that the optimum number of samplings in the above approximation depends on two parameters: (1) the fraction of hosts to infect, and (2) the density of susceptible hosts in the subset of the address space spanned by the optimal target set.

#### 3.4.1 Numerical Results

In Fig. 2 we show the results for the number of samplings versus the final fraction of infected hosts using the exact expression, and approximation (5). The results demonstrate the excellent agreement of our approximation with the exact solution over a wide range of values for the final fraction of infected hosts.

# References


[1] Milan Vojnović, Varun Gupta, Thomas Karagiannis, and Christos Gkantsidis. Sampling strategies for epidemic-style information dissemination. In *Proc. of IEEE INFOCOM*, pages 1678–1686, 2008.

[2] M. Vojnović, V. Gupta, T. Karagiannis, and C. Gkantsidis. Sampling strategies for epidemic-style information dissemination. Technical Report MSR-TR-2007-82, Microsoft Research, 2007.

[3] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. The MIT Press, 2nd edition, 2001.


# A Alternate Proof of Proposition 2

Let $J'$ and $B_{J'}(\vec{d})$ be fixed and denote $b := B_{J'}(\vec{d})$. Under the latter two constraints, maximizing $f(\vec{d})$ amounts to solving the following problem

$$\begin{array}{ll}
\text{MAX-SAM}(b): & \\
\text{maximize} & \sum_{k=1}^{J'} \omega_{(k)} \log d_{(k)} \\
\text{over} & d_{(k)} \in \left(b, \frac{n_{(k)}}{\omega_{(k)}}\right],\ k = 1, \dots, J' \\
\text{subject to} & \sum_{k=1}^{J'} \omega_{(k)} d_{(k)} = s(0) - s^0 + \sum_{k=1}^{J'} \omega_{(k)} b.
\end{array}$$

**Lemma 1.** *The solution of* MAX-SAM$(b)$ *is given by*

$$d_{(k)} = \min\left[\mu(b), \frac{n_{(k)}}{\omega_{(k)}}\right],\ k = 1, \dots, J' \qquad (6)$$

*where* $\mu(b) > b$ *is the solution of the following equation*

$$\sum_{k=1}^{J'} \omega_{(k)} \min\left[\mu(b), \frac{n_{(k)}}{\omega_{(k)}}\right] = s(0) - s^0 + \sum_{k=1}^{J'} \omega_{(k)} b. \qquad (7)$$

*Proof.* Let $x_k := d_{(k)} - b$. Problem MAX-SAM$(b)$ can be rewritten as

$$\begin{array}{ll}
\text{MAX-SAM}'(b): & \\
\text{maximize} & \sum_{k=1}^{J'} \omega_{(k)} \log(b + x_k) \\
\text{over} & x_k \in \left(0, \frac{n_{(k)}}{\omega_{(k)}} - b\right],\ k = 1, \dots, J' \\
\text{subject to} & \sum_{k=1}^{J'} \omega_{(k)} x_k = s(0) - s^0.
\end{array}$$

Let us define the Lagrangian function

$$\Lambda(x, \lambda) = \sum_{k=1}^{J'} \omega_{(k)} \log(b + x_k) - \lambda \left( \sum_{k=1}^{J'} \omega_{(k)} x_k - (s(0) - s^0) \right).$$

For fixed $\lambda > 0$ we first solve

$$\begin{array}{ll} \text{maximize} & \Lambda(x,\lambda) \\ \text{over} & x_k \in \left(0, \frac{n_{(k)}}{\omega_{(k)}} - b\right], \; k = 1,\ldots,J'. \end{array}$$

Note that the latter problem separates into $J'$ problems, for each $k = 1,\ldots,J'$,

$$\begin{array}{ll} \text{maximize} & \log(b + x_k) - \lambda x_k \\ \text{over} & x_k \in \left(0, \frac{n_{(k)}}{\omega_{(k)}} - b\right]. \end{array}$$

It is easily checked that the solution of the above problem is

$$x_k = \min\left[\frac{1}{\lambda}, \frac{n_{(k)}}{\omega_{(k)}}\right] - b.$$

Plugging into $d_{(k)} = b + x_k$ and letting $\mu := 1/\lambda$, the assertion (6) follows. Eq. (7) follows from the constraint $\sum_{k=1}^{J'} \omega_{(k)} x_k = s(0) - s^0$ in MAX-SAM'$(b)$. □

We next assume $J'$ is fixed and optimize over $b$. Note that we established that MAX-SAM can be rewritten as

$$\begin{array}{ll} \text{MAX-SAM}: & \\ \text{maximize} & g(b) \\ \text{over} & b \in B \end{array}$$

where

$$g(b) = \sum_{k=1}^{J'} \omega_{(k)} \log \frac{\min\left[\mu(b), \frac{n_{(k)}}{\omega_{(k)}}\right]}{b}$$

and

$$B = \left[\frac{\sum_{k=1}^{J'} n_{(k)} - i^0}{\sum_{k=1}^{J'} \omega_{(k)}}, \frac{\sum_{k=1}^{J'} n_{(k)} - (i^0 - i(0))}{\sum_{k=1}^{J'} \omega_{(k)}}\right]$$

and $\mu(b)$ is given by (7).

**Lemma 2.** *Function $g(b)$ is decreasing over $b \in B$.*

*Proof.* Note that we have

$$g'(b) = \sum_{k=1}^{J'} \omega_{(k)} \left(\frac{\mu'(b)}{\mu(b)} 1_{\mu(b) < n_{(k)}/\omega_{(k)}} - \frac{1}{b}\right).$$

From (7) note

$$\mu'(b) = \frac{\sum_{k=1}^{J'} \omega_{(k)}}{\sum_{k=1}^{J'} \omega_{(k)} 1_{\mu(b) < n_{(k)}/\omega_{(k)}}}.$$

Hence

$$g'(b) = \left(\sum_{k=1}^{J'} \omega_{(k)}\right)\left(\frac{1}{\mu(b)} - \frac{1}{b}\right).$$

From (7) we have that $\mu(b) > b$, and thus $g'(b) < 0$. The result follows. □

Finally, note that

$$b = \frac{\sum_{k=1}^{J'} n_{(k)} - i^0 + \sum_{k=J'+1}^{J} i_k(0)}{\sum_{k=1}^{J'} \omega_{(k)}}. \tag{8}$$

Hence, under the condition that $J'$ is fixed we have that decreasing $b$ is equivalent to relocating the infected hosts from subnets $\{J'+1, \ldots, J\}$ to subnets $\{1, \ldots, J'\}$.