國立臺灣大學電機資訊學院資訊工程學研究所

碩士論文

Department of Computer Science and Information Engineering
College of Electrical Engineering and Computer Science
National Taiwan University

Master Thesis

平面網路上防疫問題之近似演算法

An Approximation Algorithm for the Inoculation Problem for Planar Networks

邱冠凱
Kuan-Kai Chiu

指導教授：呂學一 博士
Advisor: Hsueh-I Lu, Ph.D.

中華民國 97 年 7 月
July, 2008

誌　　謝

這篇論文之所以能夠完成，首先要感謝我的指導教授呂學一老師。從一開始題目的挑選、研究經驗的分享，一直到論文的寫法和架構、口試投影片的準備，老師都非常費心地指導，讓我的研究成果得以清楚完整地呈現出來。不管是面對面還是透過電子郵件，那無數次的討論，現在回想起來都還是歷歷在目。最重要的，是能夠通過老師嚴格的督導完成這篇論文，相信以後的我，不管遇到任何困難的事，都絕對不會失去信心。

感謝王大爲老師和鍾國亮老師，願意在百忙之中抽空擔任口試委員，並對論文提供許多寶貴的意見，使這篇論文更臻完善。

再來，我也很感謝這兩年來一起陪我努力的夥伴們。冠伶學姐的開導跟指點，讓我得以保持信心繼續努力做研究。國煒學長的照顧跟搞笑，讓我由衷地感激和敬佩。還有宗灝、昱豪、弘偉、柏穎、世鵬、雅斐、文良、偉揚、婕妤等實驗室的好朋友們，讓我在這跌跌撞撞走來的兩年裡，充滿許多歡笑和回憶。

最後想要感謝的，是我的父母還有女朋友黃韻倩小姐，對於他們無怨無悔地關心和支持，這篇論文對我而言，是目前能夠回報他們最好的方式。

能不能完成一件事情，全端視自己有多渴望成功，這是我在碩士生涯裡學會最重要的一件事。僅以這篇論文，獻給所有在我拼了命努力研究的兩年裡，一直鼓勵和支持我的人們。

2008年7月

邱冠凱

平面網路上防疫問題之近似演算法

研究生：邱冠凱　　　　指導教授：呂學一 博士

國立臺灣大學電機資訊學院資訊工程學研究所

摘　要

給兩個數字 c 和 k，以及 n 個點的一般網路 G，防疫問題的定義爲：求出由 G 當中最多 k 個點所組成的點集合 S，使得 $c \cdot m + \frac{1}{n}\sum_i n_i^2$ 的值爲最小，其中 m 爲 S 的點個數，n_i 爲 $G \setminus S$ 的第 i 個連通單元。對於這個 NP-完備的問題，目前已知最好的結果是由 Aspnes、Chang 和 Yampolskiy 提出的 $O(\log^{1.5} n)$ 倍比率的近似演算法。我們在本篇論文中證明當 G 爲平面網路時，這個問題仍是 NP-完備，並且提出一個 $O(\log n)$ 倍比率的近似演算法。

An Approximation Algorithm for the Inoculation Problem for Planar Networks

Student : Kuan-Kai Chiu Advisor : Hsueh-I Lu, Ph.D.

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Abstract

For numbers c and k and an n-node graph G, the inoculation problem is to compute an S consisting of at most k nodes of G such that $c \cdot m + \frac{1}{n} \sum_i n_i^2$ is minimized, where m is the cardinality of S and n_i is the number of nodes in the i-th connected component of $G \setminus S$. The best previously known result, due to Aspnes, Chang, and Yampolskiy, for this NP-complete problem is an $O(\log^{1.5} n)$-approximation algorithm. In the present article, we focus on the special case that G is planar: We show that the problem remains NP-complete and give an $O(\log n)$-approximation algorithm for the problem.

Contents

List of Figures

Chapter 1

Introduction

For any set S, let $|S|$ denote the cardinality of S. For any node subset S of G, let $G \setminus S$ denote the graph obtained from G by deleting the nodes in S and the edges incident to the nodes in S. For any node subset S of G, let

$$\phi(S) = \sum_i n_i^2,$$

where n_i is the number of nodes in the i-th connected component of $G \setminus S$. See Figure 1.1 for an example.

Given an n-node graph G and two numbers c and k, the *inoculation problem* is to find a node subset S of G with $|S| \leq k$ that minimizes

$$c \cdot |S| + \frac{1}{n} \cdot \phi(S).$$

To address a game-theoretical model of network security (see, e.g., [2, 3, 6–8, 10–18, 20]), Aspnes, Chang, and Yampolskiy [4] formulated the problem with $k = n$ to describe the following scenario of virus attack. Suppose that each infected node incurs 1 unit of penalty and it takes c units of cost to secure a node by, say, installing an anti-virus

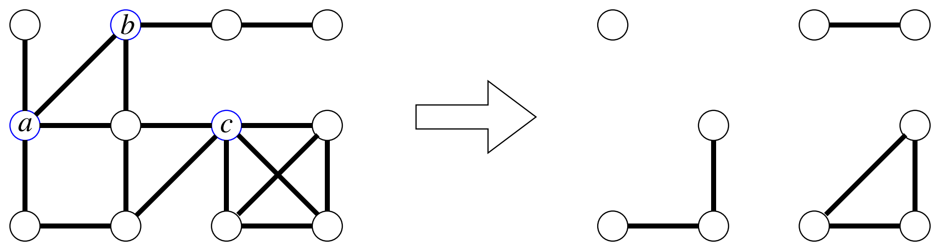

Figure 1.1: If G is the graph as shown in the left and $S = \{a, b, c\}$, then $\phi(S) = 23$.

software. The virus spreads by infecting some initial node chosen uniformly at random. An insecure node gets infected if any of its neighbors in the graph is infected. If S consists of the secured nodes, then $c \cdot |S|$ is the inoculation cost and $\frac{1}{n} \cdot \phi(S)$ is the expected penalty incurred by all infected nodes.

Aspnes et al. [4] ensured the NP-completeness of the problem. They also showed that it takes in $\tilde{O}(n^4)$ time[1] to compute an $O(\log^2 n)$-approximation solution for the problem with $k = n$. In the journal version [5], they further reduced the approximation ratio to $O(\log^{1.5} n)$ while raised the time complexity $\tilde{O}(n^{6.5})$. Moscibroda, Schimid, and Wattenhofer [19] studied the problem for highly regular and low-dimensional G.

The present article focuses on planar G, for which case we show that the problem remains NP-hard and obtain an approximation algorithm as summarized in the following theorem.

Theorem 1.1. *For any n-node planar graph G, any number c, and any positive constant ϵ, it takes $\tilde{O}(n^{6+\epsilon})$ time to compute an $O(\log n)$-approximate solution for the inoculation problem with $k = n$.*

Similar to the method of Aspnes, Chang, and Yampolskiy [4, 5], our approach repeat-

[1] The $\tilde{O}(\cdot)$ notation suppresses the polylog(n) factors.

edly removes a near-optimal sparse vertex cut from some connected component of the current graph. The near-optimal sparse vertex cut causes a sufficient decrease of risk per removed node, quantified as its cost effectiveness. We compute a node set with sufficient cost effectiveness by resorting to the approximation algorithm of Amir, Krauthgamer, and Rao [1] that finds a near optimal solution for the minimum quotient vertex cut problem for planar graphs.

The rest of the paper is organized as follows. Chapter 2 gives the preliminaries. Chapter 3 gives our algorithm. Chapter 4 concludes the paper.

Chapter 2

Preliminaries

For the rest of the paper, let G be the input n-node planar graph.

2.1 Hardness

The following lemma ensures that the inoculation problem with $k = n$ remains NP-hard even if G is planar.

Lemma 2.1. *The inoculation problem for an n-node planar graph G with $k = n$ and $\frac{1}{n} < c < \frac{3}{n}$ is NP-hard.*

Proof. The proof is modified from that of Aspnes et al. [5, Theorems 3 and 10]. Since computing the minimum cardinality of vertex cover for planar G is NP-complete [9], it suffices to show that the cardinality of any optimal solution for the inoculation problem with $k = n$ and $\frac{1}{n} < c < \frac{3}{n}$ for G is the same as the cardinality of any minimum vertex cover of G. S is a minimum vertex cover for G if and optimal solution of the inoculation problem with $c = \frac{2}{n}$.

Let S_1 be a minimum vertex cover for G. Let S_2 be an optimal solution for the inoculation problem for G with $k = n$ and $\frac{1}{n} < c < \frac{3}{n}$. Aspnes et al. [5, Theorem 10] guarantees that S_2 is a vertex cover of G. Therefore,

$$|S_1| \leq |S_2|. \tag{2.1}$$

Since S_1 and S_2 are both vertex covers of G, we have

$$\phi(S_1) = n - |S_1| \tag{2.2}$$

$$\phi(S_2) = n - |S_2|. \tag{2.3}$$

Since S_1 is a feasible solution for the inoculation problem, we have

$$c \cdot |S_2| + \frac{\phi(S_2)}{n} \leq c \cdot |S_1| + \frac{\phi(S_1)}{n}.$$

It follows from $c > \frac{1}{n}$ and Equations (2.2) and (2.3) that

$$|S_2| \leq |S_1|. \tag{2.4}$$

Combining Inequalities (2.1) and (2.4), we have $|S_1| = |S_2|$. The lemma is proved. □

2.2 Minimum quotient cut

$\langle A, B, R\rangle$ is a *cut* of a graph H if A, B, and R with $|A| \leq |B|$ form a partition of the nodes of H such that H does not have any edge with one endpoint in A and the other endpoint in B. The *quotient* of $\langle A, B, R\rangle$ is

$$\theta_H(A, B, R) = \frac{|R|}{|A| + |R|}.$$

The *minimum quotient cut problem* for H is to find a cut of H with minimum quotient.

Lemma 2.2 (Amir et al. [1])**.** *For any n-node planar graph H and any positive constant ϵ, it takes $O(n^{4+\epsilon})$ time to compute a $\frac{4}{3}(1+\frac{1}{\epsilon}+o(1))$-approximate solution for the minimum quotient cut problem for H.*

Chapter 3

Our algorithm

3.1 A reduction

For numbers $\alpha, \beta \geq 1$, a node subset $\mathbb{S}$ of G is a bicriterion (α, β)-*approximate* solution for the inoculation problem for G if the following conditions hold for $\mathbb{S}$, α, and β:

- *Condition B1*: $|\mathbb{S}| \leq \alpha \cdot k$.
- *Condition B2*: $\phi(\mathbb{S}) \leq \beta \cdot \phi(\hat{S})$ holds for any node subset $\hat{S}$ of G with $|\hat{S}| \leq k$.

Therefore, a β-approximate solution for the problem is a $(1, \beta)$-approximate solution for the problem. Our result is based on the following reduction of Aspnes et al. [5], which ensures that it suffices to focus on finding an (α, β)-approximate solution for the problem with $c = 0$.

Lemma 3.1 (Aspnes et al. [5, Corollary 13]). *If an (α, β)-approximate solution for the inoculation problem with $c = 0$ can be computed in $O(f(n))$ time, then it takes $O(n \cdot f(n))$*

time to compute a $\max(\alpha, \beta)$*-approximate solution for the inoculation problem with* $k = n$.

3.2 Finding a bicriterion approximate solution

Given a number $\gamma \geq 1$, we show how to compute an $(O(\gamma \cdot \log n), O(\gamma))$-approximate solution for the inoculation problem with $c = 0$. Observe that if $k \geq n$, then the node set of G is a trivial $(O(\gamma \cdot \log n), O(\gamma))$-approximate solution. Therefore, the rest of the section assumes $k < n$. For any node subsets S and R of G, the *cost effectiveness* of R with respect to S is

$$\mu(R, S) = \frac{\phi(S) - \phi(S \cup R)}{|R|}.$$

With the following conditions, Algorithm 1 gives the main procedure.

- *Condition C1*: $\mu(R, S) < \frac{1}{(\gamma+1)k} \cdot \phi(S)$.
- *Condition C2*: $|S \cup R| \geq \min\{n, (1 + (\gamma + 1) \log_2(n - k)) \cdot k\}$.
- *Condition C3*: $\mu(R, S) \geq \frac{1}{\gamma} \cdot \mu(\hat{S}, S)$ holds for any node subset $\hat{S}$ of G with $|\hat{S}| \leq k$.

Observe that Algorithm 1 can only abnormally abort at the step of finding a node subset R of the current $G \setminus S$ such that Condition C3 holds. If Algorithm 1 does not abnormally abort, then the following lemma ensures the correctness of Algorithm 1.

Lemma 3.2. *If Algorithm 1 does not abnormally abort, then it takes Algorithm 1 at most* n *iterations to output an* $(O(\gamma \cdot \log n), O(\gamma))$*-approximate solution for the inoculation problem with* $c = 0$.

Algorithm 1 main procedure

Let S initially be an arbitrary node subset of G with $|S| = k$. The algorithm proceeds in iterations, each of which computes a nonempty node subset R of the current $G \setminus S$ such that Condition C3 holds. If neither of Conditions C1 and C2 hold for the current S and R, then the algorithm lets $S = S \cup R$ and proceeds to the next iteration. Otherwise, the algorithm halts. The output depends on Condition C1 for the final S and R:

- If Condition C1 holds, then the algorithm outputs the final S.
- If Condition C1 does not hold, then the algorithm outputs the final $S \cup R$.

Proof. Observe that each iteration of Algorithm 1 increases the size of S by at least one. By definition of Condition C2, Algorithm 1 halts in at most n iterations. The rest of the proof argues that Algorithm 1 computes an $(O(\gamma \cdot \log n), O(\gamma))$-approximate solution.

Consider the final S and R at the last iteration of Algorithm 1. We first show that

$$|S| = O(\gamma \cdot \log n) \cdot k, \tag{3.1}$$

which holds trivially if Algorithm 1 runs for exactly one iteration. If Algorithm 1 runs for at least two iterations, let S' (respectively, R') be the set S (respectively, R) at the second-to-last iteration. Since Condition C2 does not hold for S' and R', we obtain Equation (3.1) as follows.

$$|S| = |S' \cup R'| < \min\{n, (1 + (\gamma + 1) \log_2(n - k)) \cdot k\} = O(\gamma \cdot \log n) \cdot k.$$

The following case analysis is according to whether Condition C1 holds for S and R.

Case 1: Condition C1 holds for the final S and R. We prove the lemma by showing that Conditions B1 and B2 hold with $\mathbb{S} = S$, $\alpha = O(\gamma \cdot \log n)$, and $\beta = \gamma + 1$. Note that

Condition B1 is immediate from Equation (3.1). To see that $\phi(S) \le (\gamma+1)\cdot\phi(\hat{S})$ holds for any node subset $\hat{S}$ of G with $|\hat{S}| \le k$, observe that Conditions C1 and C3 hold for S and R. Thus,

$$\frac{\phi(S)}{(\gamma+1)k} > \mu(R,S) \ge \frac{\mu(\hat{S},S)}{\gamma} = \frac{\phi(S)-\phi(S\cup\hat{S})}{\gamma\cdot|\hat{S}|} \ge \frac{\phi(S)-\phi(\hat{S})}{\gamma\cdot k}.$$

Therefore, we have Condition B2.

Case 2: Condition C1 does not hold for the final S and R. By definition of Algorithm 1, Condition C2 holds for S and R. We prove the lemma by showing that Conditions B1 and B2 hold with $\mathbb{S} = S \cup R$, $\alpha = O(\gamma\cdot\log n)$, and $\beta = 1$. We first prove Condition B1, i.e., $|S\cup R| = O(\gamma\cdot\log n)\cdot k$. By Equation (3.1), it remains to ensure $|R| \le (\gamma+1)\cdot k$ as follows

$$\frac{\phi(S)}{|R|} \ge \frac{\phi(S)-\phi(S\cup R)}{|R|} = \mu(R,S) \ge \frac{\phi(S)}{(\gamma+1)k},$$

where the last inequality is by the assumption that Condition C1 does not hold for S and R.

We next prove Condition B2, i.e., $\phi(S\cup R) \le \phi(\hat{S})$ holds for any node subset $\hat{S}$ of G with $|\hat{S}| \le k$. Let S_i (respectively, R_i) be the set S (respectively, R) at the i-th iteration of Algorithm 1. Let δ be the number of iterations executed by Algorithm 1. For notational brevity, we define $S_{\delta+1}$ to be the final $S\cup R$. For each $i = 1, \ldots, \delta$, one can verify that $S_{i+1} = S_i \cup R_i$, we have

$$\mu(R_i,S_i) = \frac{\phi(S_i)-\phi(S_i\cup R_i)}{|R_i|} = \frac{\phi(S_i)-\phi(S_{i+1})}{|S_{i+1}|-|S_i|}.$$

Since Condition C1 never holds throughout the execution, we have

$$\frac{\phi(S_i)-\phi(S_{i+1})}{|S_{i+1}|-|S_i|} \ge \frac{\phi(S_i)}{(\gamma+1)k}.$$

Thus,

$$\phi(S_{i+1}) \leq \left(1 - \frac{|S_{i+1}| - |S_i|}{(\gamma+1)k}\right)\phi(S_i). \tag{3.2}$$

Note that Condition B2 holds trivially, if Condition C2 holds with $|S \cup R| = n$. For the rest of the proof, we have Condition C2 holds with

$$|S \cup R| \geq (1 + (\gamma+1)\log_2(n-k)) \cdot k.$$

Since it may not disconnect G by removing S_1 from G with $|S_1| = k$, we have $\phi(S_1) \leq (n-k)^2$. Since it may divide G into $n - |\hat{S}|$ nodes by removing $\hat{S}$ from G with $|\hat{S}| \leq k$, we have $\phi(\hat{S}) \geq n-k$. Thus, we have the following result.

$$\begin{aligned}
\phi(S \cup R) &= \phi(S_{\delta+1}) \\
&\leq \phi(S_1) \cdot \prod_{i=1}^{\delta}\left(1 - \frac{|S_{i+1}| - |S_i|}{(\gamma+1)k}\right) \\
&\leq (n-k)^2 \cdot \prod_{i=1}^{\delta}\left(1 - \frac{|S_{i+1}| - |S_i|}{(\gamma+1)k}\right) \\
&\leq (n-k)^2 \cdot \prod_{i=1}^{\delta}\left(1 - \frac{1}{(\gamma+1)k}\right)^{|S_{i+1}|-|S_i|} \\
&= (n-k)^2 \cdot \left(1 - \frac{1}{(\gamma+1)k}\right)^{|S_{\delta+1}|-|S_1|} \\
&\leq (n-k)^2 \cdot \left(1 - \frac{1}{(\gamma+1)k}\right)^{(\gamma+1)k\log_2(n-k)} \\
&< (n-k)^2 \cdot \left(\frac{1}{2}\right)^{\log_2(n-k)} \\
&= (n-k) \\
&\leq \phi(\hat{S}),
\end{aligned}$$

where the first inequality is by Inequality (3.2). Therefore, we have Condition B2. □

3.3 Finding a node set with good cost effectiveness

The following lemma ensures the feasiblity of each iteration of Algorithm 1.

Lemma 3.3. *For any positive constant ϵ, if $\gamma = 8(2 + \frac{1}{\epsilon})$, then each iteration of Algorithm 1 requires $O(n^{4+\epsilon})$ time and does not abnormally abort.*

Proof. Clearly, it suffices to focus on the step to compute a node subset R of the current $G \setminus S$ such that Condition C3 holds. At first, for each connected component H of $G \setminus S$, we compute a node subset $\mathbb{R}$ of H such that the following condition holds.

- *Condition D1*: $\mu(\mathbb{R}, S) \geq \frac{1}{\gamma} \cdot \mu(\hat{R}, S)$ holds for any node subset $\hat{R}$ of H.

The detail is left to the second part. Then, by choosing the one with maximum $\mu(\mathbb{R}, S)$ over all $\mathbb{R}$, we can derive the node subset R such that Condition D1 holds with $\mathbb{R} = R$ for any H of $G \setminus S$. Let V_H be the node subset of H. By definition of ϕ, we have

$$\phi(S) - \phi(\hat{S}) = \sum_{H \in G \setminus S} \left(\phi(S) - \phi(S \cup \{\hat{S} \cap V_H\}) \right),$$

then

$$\begin{aligned}
\mu(\hat{S}, S) &= \frac{\phi(S) - \phi(S \cup \hat{S})}{|\hat{S}|} \\
&= \frac{\sum_{H \in G \setminus S} \left(\phi(S) - \phi(S \cup \{\hat{S} \cap V_H\}) \right)}{|\hat{S}|} \\
&= \frac{\sum_{H \in G \setminus S} \left(\mu(\hat{S} \cap V_H, S) \cdot |\hat{S} \cap V_H| \right)}{|\hat{S}|}.
\end{aligned}$$

Since $\hat{S} \cap V_H$ is a node subset of H, and Condition D1 holds with $\mathbb{R} = R$ for any H of

$G \setminus S$, Condition C3 holds by the follows

$$\begin{aligned}\mu(\hat{S}, S) &= \frac{\sum_{H \in G \setminus S} \left(\mu(\hat{S} \cap V_H, S) \cdot |\hat{S} \cap V_H|\right)}{|\hat{S}|} \\ &\leq \frac{\sum_{H \in G \setminus S} \left(\gamma \cdot \mu(R, S) \cdot |\hat{S} \cap V_H|\right)}{|\hat{S}|} \\ &\leq \gamma \cdot \mu(R, S).\end{aligned}$$

The rest of the proof proves that, for any connected component H of $G \setminus S$, it takes $O(|V_H|^{4+\epsilon})$ time to compute a node subset $\mathbb{R}$ of H such that Condition D1 holds. Therefore, we can derive R in $O(n^{4+\epsilon})$ time, which proves Lemma 3.3.

For any node subset $\hat{R}$ of H, let $\langle \hat{A}, \hat{B}, \hat{R} \rangle$ be a cut of H, and $|\hat{A}|$ is maximized with respect to $\hat{R}$. We give the proof in three parts.

1. $\mu(\hat{R}, S) \leq \frac{3|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})}$.
2. A cut $\langle \mathbb{A}, \mathbb{B}, \mathbb{R} \rangle$ with $\theta_H(\mathbb{A}, \mathbb{B}, \mathbb{R}) \leq \frac{4\left(2+\frac{1}{\epsilon}\right)}{3} \cdot \theta_H(\hat{A}, \hat{B}, \hat{R})$ can be computed in $O(|V_H|^{4+\epsilon})$ time.
3. $\theta_H(\mathbb{A}, \mathbb{B}, \mathbb{R}) \geq \frac{|V_H|}{2 \cdot \mu(\mathbb{R}, S)}$.

Combining the three statements, we can compute a node subset $\mathbb{R}$ of H in $O(|V_H|^{4+\epsilon})$ time such that

$$\mu(\hat{R}, S) \leq \frac{3|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})} \leq \frac{4\left(2+\frac{1}{\epsilon}\right)|V_H|}{\theta_H(\mathbb{A}, \mathbb{B}, \mathbb{R})} \leq 8\left(2+\frac{1}{\epsilon}\right) \cdot \mu(\mathbb{R}, S) = \gamma \cdot \mu(\mathbb{R}, S).$$

Therefore, we have Condition D1.

Let V_H^i be the node set of the i-th connected component of the remaining graph induced by removing $\hat{R}$ from H. By definition of ϕ, we have

$$\phi(S) - \phi(S \cup \hat{R}) = |V_H|^2 - \sum_i |V_H^i|^2. \tag{3.3}$$

Since $\mathbb{R}$ is a node subset of H, Equation (3.3) also holds with $\hat{R} = \mathbb{R}$ for the rest of the proof.

The first statement can be proved by two cases of $|\hat{B}|$.

- $|\hat{B}| \le \frac{2|V_H|}{3}$. From Equation (3.3), we have

$$\begin{aligned}
\mu(\hat{R}, S) &= \frac{\phi(S) - \phi(S \cup \hat{R})}{|\hat{R}|} \\
&= \frac{|V_H|^2 - \sum_i |V_H^i|^2}{|\hat{R}|} \\
&\le \frac{|V_H|^2 - \sum_i |V_H^i|}{|\hat{R}|} \\
&= \frac{|V_H|^2 - (|\hat{A}| + |\hat{B}|)}{|\hat{R}|} \\
&= \frac{|V_H|^2 - (|V_H| - |\hat{R}|)}{|\hat{R}|} \\
&= \frac{|V_H|(|V_H| - 1)}{|\hat{R}|} + 1 \\
&= \frac{|V_H|(|\hat{A}| + |\hat{B}| + |\hat{R}| - 1)}{|\hat{R}|} + 1 \\
&= \frac{|V_H|(|\hat{A}| + |\hat{R}|)}{|\hat{R}|} + \frac{|V_H|(|\hat{B}| - 1)}{|\hat{R}|} + 1 \\
&\le \frac{|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})} + \frac{2|V_H|(|V_H| - |\hat{B}|)}{|\hat{R}|} \\
&= \frac{|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})} + \frac{2|V_H|(|\hat{A}| + |\hat{R}|)}{|\hat{R}|} \\
&= \frac{|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})} + \frac{2|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})} \\
&= \frac{3|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})}.
\end{aligned}$$

- $|\hat{B}| > \frac{2|V_H|}{3}$. We prove that there is only one connected component in $\hat{B}$. Assume for contradiction that $\hat{B}$ consists of more than one connected components. The following two cases show that there exists a cut $\langle \mathcal{A}', \mathcal{B}', \hat{R} \rangle$ with $|\mathcal{A}'| > |\hat{A}|$, contradicting to that $|\hat{A}|$ is maximized with respect to $\hat{R}$.

- If there exists a connected component H_b of $\hat{B}$ with $|H_b| \le \frac{|V_H|}{3}$, we let $\mathcal{A}' = \min\{\hat{A} \cup H_b, \hat{B} \setminus H_b\}$, then $|\mathcal{A}'| > |\hat{A}|$.
- Otherwise, we let $\mathcal{A}'$ be any connected component of $\hat{B}$, then $|\mathcal{A}'| > \frac{|V_H|}{3} > |\hat{A}|$.

Thus, we have

$$\begin{aligned}
\mu(\hat{R}, S) &= \frac{\phi(S) - \phi(S \cup \hat{R})}{|\hat{R}|} \\
&= \frac{|V_H|^2 - \sum_i |V_H^i|^2}{|\hat{R}|} \\
&= \frac{|V_H|^2 - |\hat{B}|^2 - \sum_{V_H^i \in \hat{A}} |V_H^i|^2}{|\hat{R}|} \\
&\le \frac{|V_H|^2 - |\hat{B}|^2 - |\hat{A}|}{|\hat{R}|} \\
&= \frac{|V_H|^2 - (|V_H| - |\hat{A}| - |\hat{R}|)^2 - |\hat{A}|}{|\hat{R}|} \\
&= \frac{2|V_H|(|\hat{A}| + |\hat{R}|) - (|\hat{A}| + |\hat{R}|)^2 - |\hat{A}|}{|\hat{R}|} \\
&< \frac{2|V_H|(|\hat{A}| + |\hat{R}|)}{|\hat{R}|} \\
&= \frac{2|V_H|}{\theta_H(\hat{A}, \hat{B}, \hat{R})},
\end{aligned}$$

where the second equality is by Equation (3.3).

The second statement follows immediately from Lemma 2.2. We then prove the third statement. Since $|\mathbb{A}| \le |\mathbb{B}|$, we have

$$|\mathbb{B}| \ge \frac{|\mathbb{A}| + |\mathbb{B}|}{2} = \frac{|V_H| - |\mathbb{R}|}{2}. \tag{3.4}$$

From Equation (3.3) and Inequality (3.4), we have

$$
\begin{aligned}
\mu(\mathbb{R}, S) &= \frac{\phi(S) - \phi(S \cup \mathbb{R})}{|\mathbb{R}|} \\
&= \frac{|V_H|^2 - \sum_i |V_H^i|^2}{|\mathbb{R}|} \\
&\geq \frac{|V_H|^2 - |\mathbb{A}|^2 - |\mathbb{B}|^2}{|\mathbb{R}|} \\
&= \frac{(|\mathbb{A}| + |\mathbb{B}| + |\mathbb{R}|)^2 - |\mathbb{A}|^2 - |\mathbb{B}|^2}{|\mathbb{R}|} \\
&= \frac{2|\mathbb{A}||\mathbb{B}| + 2|V_H||\mathbb{R}| - |\mathbb{R}|^2}{|\mathbb{R}|} \\
&\geq \frac{2|\mathbb{A}||\mathbb{B}| + |V_H||\mathbb{R}|}{|\mathbb{R}|} \\
&\geq \frac{|\mathbb{A}|(|V_H| - |\mathbb{R}|) + |V_H||\mathbb{R}|}{|\mathbb{R}|} \\
&= \frac{|V_H|(|\mathbb{A}| + |\mathbb{R}|)}{|\mathbb{R}|} - \frac{|\mathbb{A}||\mathbb{R}|}{|\mathbb{R}|} \\
&\geq \frac{|V_H|}{\theta_H(\mathbb{A}, \mathbb{B}, \mathbb{R})} - \frac{|V_H|(|\mathbb{A}| + |\mathbb{R}|)}{2|\mathbb{R}|} \\
&= \frac{|V_H|}{2 \cdot \theta_H(\mathbb{A}, \mathbb{B}, \mathbb{R})}.
\end{aligned}
$$

□

3.4 Proving Theorem 1.1

Proof. Lemmas 3.2 and 3.3 together ensure that an $(O(\log n), O(1))$-approximate solution for the inoculation problem with $c = 0$ can be found in $O(n^{5+\epsilon})$ time. The theorem follows immediately from Lemma 3.1. □

Chapter 4

Concluding remarks

We leave open the approximability of the inoculation problem for general parameter k. It would be interesting to see if our techniques can be extended to work for this general version for planar graphs.

Bibliography

[1] E. Amir, R. Krauthgamer, and S. Rao. Constant factor approximation of vertex-cuts in planar graphs. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, pages 90–99, New York, NY, USA, 2003. ACM.

[2] R. Anderson and T. Moore. The economics of information security. *Science Magazine*, 314(5799):610–613, October 2006.

[3] R. Anderson and T. Moore. Information security economics - and beyond. In *Proceedings of the 27th Annual International Cryptology Conference*, volume 4622 of *Lecture Notes in Computer Science*, pages 68–91. Springer, 2007.

[4] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 43–52, Vancouver, British Columbia, Canada, January 23-25 2005.

[5] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *Journal of Computer and System Sciences*, 72(6):1077–1093, 2006.

[6] N. Chen. On the approximability of influence in social networks. In *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1029–1037, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[7] R. Cole, Y. Dodis, and T. Roughgarden. Bottleneck links, variable demand, and the tragedy of the commons. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithm*, pages 668–677, New York, NY, USA, 2006. ACM.

[8] R. Eidenbenz, Y. A. Oswald, S. Schmid, and R. Wattenhofer. Mechanism design by creditability. In *Proceedings of the 1st International Conference on Combinatorial Optimization and Applications*, volume 4616 of *Lecture Notes in Computer Science*, pages 208–219. Springer, 2007.

[9] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete problems. In *Proceedings of the 6th Annual ACM Symposium on Theory of Computing*, pages 47–63, New York, NY, USA, 1974. ACM.

[10] M. Gelastou, M. Mavronicolas, V. G. Papadopoulou, A. Philippou, and P. G. Spirakis. The power of the defender. In *Proceedings of the 26th IEEE International Conference on Distributed Computing System Workshops*, page 37. IEEE Computer Society, 2006.

[11] G. Giakkoupis, A. Gionis, E. Terzi, and P. Tsaparas. Models and algorithms for network immunization.

[12] M. Mavronicolas, V. G. Papadopoulou, G. Persiano, A. Philippou, and P. G. Spirakis. The price of defence and fractional matchings. In *Proceedings of the 8th*

International Conference on Distributed Computing and Networking, volume 4308 of *Lecture Notes in Computer Science*, pages 115–126. Springer, 2006.

[13] M. Mavronicolas, V. G. Papadopoulou, A. Philippou, and P. G. Spirakis. A graph-theoretic network security game. In *Proceedings of the 1st International Workshop on Internet and Network Economics*, volume 3828 of *Lecture Notes in Computer Science*, pages 969–978. Springer, 2005.

[14] M. Mavronicolas, V. G. Papadopoulou, A. Philippou, and P. G. Spirakis. Network game with attacker and protector entities. In *Proceedings of the 16th Annual International Symposium on Algorithms and Computation*, volume 3827 of *Lecture Notes in Computer Science*, pages 288–297. Springer, 2005.

[15] M. Mavronicolas, V. G. Papadopoulou, A. Philippou, and P. G. Spirakis. A graph-theoretic network security game. *International Journal of Autonomous and Adaptive Communications Systems*, June 2008.

[16] M. Mavronicolas, V. G. Papadopoulou, A. Philippou, and P. G. Spirakis. A network game with attackers and a defender. *Algorithmica*, 51(3):315–341, June 2008.

[17] D. Meier, Y. A. Oswald, S. Schmid, and R. Wattenhofer. On the windfall of friendship: inoculation strategies on social networks. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 2008.

[18] T. Moscibroda. *Locality, scheduling, and selfishness: algorithmic foundations of highly decentralized networks*. PhD thesis, ETH Zurich, 2006.

[19] T. Moscibroda, S. Schmid, and R. Wattenhofer. When selfish meets evil: Byzantine players in a virus inoculation game. In *Proceedings of the 25th Annual ACM Symposium on Principles of Distributed Computing*, pages 35–44, New York, NY, USA, 2006. ACM.

[20] Z. Nikoloski. Defence strategies against network worms: formulation, evaluation, and comparison.