Journal of Graph Algorithms and Applications
http://jgaa.info/ vol. 15, no. 5, pp. 683–699 (2011)

Variants of Spreading Messages

T V Thirumala Reddy C Pandu Rangan

Department of Computer Science and Engineering,
Indian Institute of Technology Madras, Chennai

Abstract

In a distributed computing environment a faulty node could lead other nodes in the system to behave in a faulty manner. An initial set of faults could make all the nodes in the system become faulty. Such a set is called an irreversible dynamo. This is modelled as spreading a message among individuals V in a community $G = (V, E)$ where E represents the acquaintance relation. A particular individual will believe a message if some of the individual's acquaintances believe the same and forward the believed messages to its neighbours. We are interested in finding the minimum set of initial individuals to be considered as convinced, called the *min-seed*, such that every individual in the community is finally convinced. In this paper we give an upper bound on the cardinality of the *min-seed* for undirected graphs. We consider some interesting variants of the problem and analyse their complexities and give some approximate algorithms.

Submitted: April 2010	Reviewed: November 2010	Revised: February 2011	Reviewed: June 2011	Revised: July 2011
Reviewed: August 2011	Revised: August 2011	Accepted: August 2011	Final: September 2011	Published: October 2011
	Article type: Regular paper		Communicated by: M. S. Rahman and S. Fujita	

E-mail addresses: tiru114@gmail.com (T V Thirumala Reddy) prangan55@gmail.com (C Pandu Rangan)

1 Introduction

In a distributed computing environment a node could become faulty. A faulty node could make some other nodes in the system behave in a faulty manner. In order to design a fault tolerant system, we need to examine some faulty nodes as well as the cumulative effect of these initial faulty nodes on other nodes of the system. We are interested in the patterns of the initial faults that can occur and then could lead all the other nodes in the system behave in a faulty manner. The initial set of faults that leads all the nodes to become faulty is called a *dynamic monopoly* in the system. Faults can be temporary or permanent. If we consider the faults of the system as permanent then the problem is called as the *irreversible dynamo* [3]. This problem is modelled in graph theory as the Spreading Message problem.

In the Spreading Message problem we have a set of individuals representing the vertices in a graph and the acquaintance relation of individuals represents the edges of the graph. An individual believes a message when he receives it from his acquaintances, who are already convinced by the message. Every vertex v has a threshold $\alpha(v)$. A vertex is considered as *convinced*, if at least $\alpha(v)$ of its neighbours are already convinced. We are interested in finding a minimum cardinality set of individuals to be convinced who can eventually convince all the individuals. The cardinality of this set is called the *min-seed*. We can now observe that in the spreading messages problem, vertices represent nodes in a distributed environment and a convinced vertex represents a faulty node. The threshold function $\alpha(v)$ represents fault tolerance of an individual node. This problem was considered by Peleg [17], where there were white and black nodes corresponding to good and faulty nodes. The problem was largely studied on random graphs [19]. A variant of the problem is to consider the majority scenarios, where a vertex will be convinced if majority of its neighbours are convinced. Majority scenarios like strict majority and weak majority were considered in the past on tori [10], butterfly graph [13] and chordal rings [9]. The problem with an arbitrary threshold function was first considered by Ching-Lueh Chang and Yuh-Dauh Lyuu [4, 5], where it is shown to be NP-complete on arbitrary undirected graphs.

In our paper, we give an upper bound on *min-seed* of unbounded spreading messages of an arbitrary undirected graph and show that the problem is NP-Complete on bipartite graphs. The first variant we consider is spreading messages within one round. For this variant we give a lower bound and an $(H(\alpha_M) + H(\Delta))$ approximation algorithm, where $H(n)$ represents the sum of first n terms in the harmonic series, α_M represents the maximum threshold of vertices in the graph and Δ represents the maximum degree of a graph. We also show that this variant is APX-Complete on bounded degree 3 graphs and on p-claw free graphs we provide an $\frac{c\alpha_M \cdot (p-1)}{\alpha_m}$ approximation algorithm, where α_m represents minimum threshold of vertices in the graph and $c = \max\{1, \dfrac{\alpha_m}{p-1}\}$ and $c \in \mathbb{R}^+$. Another variant we consider is spreading messages within k rounds. We show that this variant is NP-complete on arbitrary undirected graphs. Then we introduce spreading messages problem with real thresholds and belief factors. We show that this variant is NP-Complete on cliques and complete bipartite graphs. Finally we consider spreading messages with each individual having r radius of coverage and we give an $(H(\alpha_M) + H(n))$ approximation algorithm for arbitrary undirected graphs. A preliminary version of this paper has been presented at [18].

2 Notation and Definitions

A simple graph is a collection of vertices V and edges E represented as $G = (V, E)$, where each edge is an unordered pair of distinct vertices. In this paper we consider only simple undirected connected graphs. We also assume that $|V| > 1$ for the graphs we consider. The set of neighbours of a vertex v is denoted by $N(v)$ and $N[v] = N(v) \cup \{v\}$. The distance between two vertices in a graph is the number of edges in a shortest path connecting them. $N^r(v)$ is the set of all vertices whose distance from v is less than or equal to r and $N^r[v] = N^r(v) \cup \{v\}$. For a subset $S \subseteq V$, $N_S(v) = N(v) \cap S$ and $N_S[v] = N_S(v) \cup \{v\}$. The degree of a vertex v, $d(v)$, is defined as $d(v) = |N(v)|$. The maximum degree of G, $\Delta = \max_{v \in V}\{d(v)\}$. The maximum threshold of G, $\alpha_M = \max_{v \in V}\{\alpha(v)\}$. The minimum threshold of G, $\alpha_m = \min_{v \in V}\{\alpha(v)\}$.

Let $G = (V, E)$ be a simple undirected graph. Assume V represents a set of individuals and the acquintance relation between them is represented with edges. A vertex or an individual believes a message when he receives it from his acquaintances, who are already convinced by the message. Let $\alpha : V \to \mathbb{N}$ be a threshold function such that $1 \leq \alpha(v) \leq d(v)$ for all $v \in V$, where a vertex v is *convinced* if at least $\alpha(v)$ of neighbours of v are already convinced. We are interested in finding a minimum cardinality set of individuals to be convinced who can eventually convince all the individuals. The cardinality of this set is called the *min-seed*.

Definition 1 *Let $S_0 \subseteq V$ be a vertex subset. Then spreading of a message will happen in rounds. Let $C_0 = S_0 \subseteq V$ be the initial set of vertices considered directly convinced.*

$$S_i = \{x | \alpha(x) \leq |C_{i-1} \cap N(x)|\} \quad , \quad C_i = S_i \cup C_{i-1}$$

Unbounded Spreading Messages: *S_0 is called* seed *if and only if $\bigcup_{i=0}^{\infty} S_i = V$.* MIN-SEED (G, α, ∞) *is defined as $\min_S(|S|)$ for all possible seeds S. A seed S with $|S| =$* MIN-SEED (G, α, ∞) *is called an optimum seed.*

Bounded Spreading Messages within k Rounds: *S_0 is called* seed *if and only if $\bigcup_{i=0}^{k} S_i = V$.* MIN-SEED (G, α, k) *is defined as $\min_S(|S|)$ for all seeds S.*

Unbounded Spreading Messages With Radius of Coverage: *We consider another variant of* SPREADING MESSAGES *in which we introduce the new term* radius of coverage. *If a vertex v is convinced and let its radius of coverage be r*

then v *can send message to all vertices reachable from* v *with distance less than or equal to* r*. We consider the case where all vertices have the same radius of coverage* r*. If* $r = 1$ *then this problem is the same as the original version.*
Unbounded Spreading Messages With Real Thresholds and Belief Factors*: Let* $\beta : E \to \mathbb{Q}^+$ *be a mapping such that* $0 < \beta(u,v) \leq 1$ *and* $\alpha : V \to \mathbb{Q}^+$ *be a mapping such that* $1 \leq \alpha(v) \leq \sum_{u \in N(v)} \beta(u,v)$*, for all* $v \in V$*. For an edge* (u,v)*,* $\beta(u,v)$ *is called belief factor of* (u,v)*. In any round* C *denotes the set of convinced vertices and* $N_c(v)$ *denote the set of convinced neighbours of a vertex* v*. A vertex* v *is convinced by a message if* $\sum_{u \in N_c(v)} \beta(u,v) \geq \alpha(v)$*.*
The definition of C_i *remains same but the* S_i *is defined as follows:*

$$S_i = \left\{ x | \alpha(x) \leq \sum_{u \in N(v) \cap C_{i-1}} \beta(u,v) \right\} \quad , \quad C_i = S_i \cup C_{i-1}$$

A set $S_0 \subseteq V$ *is called* seed *if and only if* $\bigcup_{i=0}^{\infty} S_i = V$*.* Min-Seed$(G, \alpha, \beta, \infty)$ *is defined as* $\min_S(|S|)$ *for all seeds* S*.*

Unified Notation: To cover all the variants, we give a unified notation for the problem. The problem Min-Seed(G, α, k, β, r):

- r - radius of coverage
- α - threshold values function
- k - number of stages, k $= \mathcal{N} \cup \{\infty\}$
- β - belief function
  if r, β values are equal to 1 then they are not written explicitely.

We denote by α^{min} the minimum threshold function, i.e., $\forall v$, $\alpha(v) = 1$.
We denote by α^{max} the maximum threshold function, i.e., $\forall v$, $\alpha(v) = d(v)$.
Seed(G, α, k, β, r) is the problem of computing any *seed*.
Min-Seed(G, α, k, β, r) is the problem of computing the minimum size *seed*.
Min-Seed-Cnt(G, α, k, β, r) is the number of vertices in the Min-Seed solution, i.e., size of the Min-Seed. Given $G, \alpha, k, \beta, r, s$, where k is a positive integer, let Min-Seed-D be the problem of deciding whether Min-Seed-Cnt$(G, \alpha, k, \beta, r) \leq s$.

Lemma 1 *From the definition of the problem, we can observe the following statements:*

1. $k \leq k' \Rightarrow$ Min-Seed-Cnt$(G, \alpha, k', \beta, r) \leq$ Min-Seed-Cnt(G, α, k, β, r)
2. $r \leq r' \Rightarrow$ Min-Seed-Cnt$(G, \alpha, k, \beta, r') \leq$ Min-Seed-Cnt(G, α, k, β, r)

3. $\beta \preceq \beta' \Rightarrow$ MIN-SEED-CNT$(G, \alpha, k, \beta', r) \leq$ MIN-SEED-CNT(G, α, k, β, r)
4. $\alpha \preceq \alpha' \Rightarrow$ MIN-SEED-CNT$(G, \alpha, k, \beta, r) \leq$ MIN-SEED-CNT$(G, \alpha', k, \beta, r)$

The relation $\beta \preceq \beta'$ specifies that $\forall e \in E, \beta(e) \leq \beta'(e)$. Similarly $\alpha \preceq \alpha'$ specifies that $\forall v \in V, \alpha(v) \leq \alpha'(v)$.

Given an instance of SEED(G, α, k, β, r) problem and set of vertices of G, by using brute-fource technique we can check whether the given set of vertices is *seed* for all the variants and their combinations in polynomial amount of time. Hence we have the following lemma.

Lemma 2 *For any given graph G,* MIN-SEED-D$(G, \alpha, k, \beta, r, s)$ *is in NP.*

Some well known $NP-complete$ problems used in this paper are:

Definition 2 *A vertex-cover of an undirected graph $G = (V, E)$ is a subset of V, say V', such that if $edge(u, v)$ is an edge of G then either $u \in V'$ or $v \in V'$ (or $u, v \in V'$).* VERTEX-COVER-OPT *is the problem of computing the minimum size vertex-cover.*

Definition 3 *[15] Given a universal set of elements U and a set S, where S is set of subsets of U. A Set-cover is a collection of the subsets in S whose union is U.* SET-COVER-OPT *is the problem of computing the minimum size set-cover.*

Definition 4 *[15] A dominating set of an undirected graph $G = (V, E)$ is a subset of V, say D, such that every vertex of $D \backslash V$ is a neighbour of at least one vertex in D.* DOMINATING-SET-OPT *is the problem of computing the minimum size dominating-set.*

Definition 5 *[11] Given $G = (V, E)$ and let r be a positive integer. A non-empty subset $D \subseteq V$ is a $r-dominating$ set if every vertex in $V \setminus D$ is within a distance r from at least one vertex of D.*

Definition 6 *[11] Given two NP optimization problems P and Q and a polynomial transformation f from instances of P to instances of Q, we say that f is an $L-reduction$ if there are positive constants a and b such that for every instance x of P*

1. $opt_Q(f(x)) \leq a \cdot opt_P(x)$,
2. *for every feasible solution y of $f(x)$ with objective value $m_Q(f(x), y) = c_2$ we can in polynomial time find a solution y' of x with $m_P(x, y') = c_1$ such that $|opt_P(x) - c_1| \leq b \cdot |opt_Q(f(x)) - c_2|$.*

To show the APX-completeness of a problem $P \in APX$, it is enough to show that there is an L-reduction from some APX-complete problem to P.

Definition 7 *[11] A graph $G = (V, E)$ is called a p-claw free graph if for all the vertices v, the subgraph induced by $N(v)$ does not have an independent set of size p.*
Alternately, a graph $G = (V, E)$ is called a p-claw free graph if there is no induced subgraph of G isomorphic to the star graph K_{1p}.

3 Complexity Results

Theorem 1 MIN-SEED-D(G, α, k, s) *is NP-complete for all k.*

Proof: The proof for this theorem comntain two cases. First case is when $k \geq 2$ and second case is when $k = 1$. The proof for each case contain different reductions from different problems. If $k \geq 2$ then we reduce SET-COVER to MIN-SEED-D. If $k = 1$ then we reduce VERTEX-COVER to MIN-SEED-D.

Case 1: In this case $k \geq 2$. Given an instance of the SET-COVER with an universal set $U = \{x_1, x_2, x_3, ..., x_n\}$, a set of subsets $S = \{S_1, S_2, ..., S_m\}$ and an integer s, where $S_i \subseteq U$. Construct a bipartite graph $G = (X, Y, E)$, with $|X| = \{s_1, s_2, s_3, ..., s_m\}$ and $|Y| = \{x_1, x_2, x_3, ..., x_n\}$. That is the vertex set X contains a vertex for every set of S and the vertex set Y contains a vertex for every element of U. If the element $x_j \in S_i$ then connect the vertex s_i to the vertex x_j. Set $\alpha(s_i) = d(s_i)$, $\forall s_i \in X$ and $\alpha(x_j) = 1$, $\forall x_j \in Y$.

Let there exist a solution for SET-COVER problem with size s. Now the solution for the SEED(G, α, k) problem is: for every set S_i in the SET COVER solution, choose the vertex s_i of X in the seed. These s vertices of X first convince all the vertices of Y. Then the remaining vertices of X get convinced. Because $\forall s_i \in X, \alpha(s_i) = d(s_i)$ and all the neighbours of s_i (vertices of Y) are convinced. Hence MIN-SEED-CNT$(G, \alpha, 2) \leq$ OPT-SET-COVER(S, U).

Let there exist a solution for SEED(G, α, k) with size s. If the vertex s_i of X is in the SEED(G, α, k) then include set S_i in set cover solution. If a vertex $x_j \in Y$ is in SEED(G, α, k) then choose any neighbour of the vertex x_j. Let say the vertex s_k of X is chosen, then include the set S_k in set cover solution. Now we prove that the sets chosen cover all the elements of U. Let us assume that some element $x_j \in U$ not covered. Consider the possibilities of how the vertex $x_j \in Y$ is convinced. Definitely the vertex x_j and the neighbours of x_j are not in SEED(G, α, k) solution, so the vertex x_j must be convinced by its neighbours. Neighbours of the vertex x_j are convinced if and only if the vertex x_j is convinced because $\forall s_i \in N(x_j), \alpha(s_i) = d(s_i)$. This implies that the vertex x_j never gets convinced, which is a contradicting statement. Therefore the sets we choose form a SET COVER solution. The graph constructed here is a bipartite graph and this proof works for $k = 2$ as well. Hence MIN-SEED-CNT$(G, \alpha, \infty) \geq$ OPT-SET-COVER(S, U).

Case 2: Given an instance of VERTEX COVER problem with a graph $G = (V, E)$ and a positive integer s, construct an instance of the SEED problem with the same graph G. Define $\alpha(v) = d(v)$ for all $v \in V$.

Let there exist a solution for VERTEX COVER problem with s vertices. These s vertices also give us a solution for SEED$(G, \alpha, 1)$, because if a vertex v is not in the VERTEX COVER then all its neighbours must be in the VERTEX COVER.

So all the vertices not in VERTEX COVER get convinced if we convince vertices in VERTEX COVER.

Let there exist a solution for SEED$(G, \alpha, 1)$ with size s. Now we prove that these s vertices gives a solution for VERTEX COVER. Let us assume that some edge (u, v) is not covered. Both vertices u and v are not in the SEED solution. Now consider the possibilities of how the vertices u and v get convinced. In order to convince the vertex u, first the vertex v must be convinced. Similarly, in order to convince the vertex v, first vertex u must be convinced. This leads to a contradiction that neither of the vertex u nor the vertex v gets convinced. So one of the vertices u, v must be in SEED$(G, \alpha, 1)$ solution. □

Corollary 1 *For every graph G and every valid threshold function α,* MIN-SEED-CNT$(G, \alpha, 1)$ *is bounded above by the* Vertex Cover Number *of G.*

By using the graph constructed in the Theorem 1 Case 1, we can prove that MIN-SEED-D$(G', \alpha, 2, \beta, s)$ is NP-complete, where G' is a complete bipartite graph formed by adding missing edges in G. Define function β like this: $\beta = 1$ for all edges belong to G and $\beta = \frac{1}{(m+n)^2}$ for all newly added edges. This makes all the newly added edges dummy, and the above proof still works to prove the following theorem.

Theorem 2 MIN-SEED-D$(G, \alpha, 2, \beta, s)$ *is NP-complete when G is a complete bipartite graph.* □

By using the graph constructed in the Theorem 1 Case 2, we can prove that MIN-SEED-D$(G', \alpha, 1, \beta, s)$ is NP-complete, where G' is a complete graph formed by adding missing edges in G. Define function β like this: $\beta = 1$ for all edges belong to G and $\beta = \frac{1}{n}$ for all newly added edges. This make all the newly added edges dummy and the above proof still work to prove the following theorem.

Theorem 3 MIN-SEED-D$(G, \alpha, 1, \beta, s)$ *is NP-complete when G is a complete graph.* □

Theorem 4 *For every $r \geq 1$,* MIN-SEED-D$(G, \alpha, 1, r, s)$ *is NP-complete.*

Proof: First we show how to reduce the $r-$ DOMINATING SET problem to the MIN-SEED-D$(G, \alpha, 1, r, s)$ problem.

Construction: Given an instance of $r-$ DOMINATING SET problem with graph $G = (V, E)$ and a positive integer s, construct an instance of MIN-SEED-D$(G, \alpha, 1, r, s)$ problem with same graph G. Define $\alpha(v) = 1$ for all $v \in V$.

Let there exist a solution for the $r-$ DOMINATING SET problem with s vertices. These s vertices also give us a *seed*, because if a vertex v is not in $r-$ DOMINATING SET then one of the vertex at distance r must be in $r-$ DOMINATING SET. So all vertices not in $r-$ DOMINATING SET get convinced if we convince vertices in $r-$ DOMINATING SET in one round.

Let there exist a solution for $\textsc{Seed}(G, \alpha, 1, r)$ of size s. Now we prove that these s vertices also give an $r-$ DOMINATING SET. Let us assume the contradiction that some vertex v is not $r-$ *dominated* by any vertex. The vertex v and the vertices in $N^r(v)$ are not in $\textsc{Seed}(G, \alpha, 1, r)$ solution. The vertices belongs to $N^r(v)$ get convinced in one round but for convincing v we need two rounds. This leads to a contradiction. So either v or a vertex from $N^r(v)$ must be in $\textsc{Seed}(G, \alpha, 1, r)$.

As we are setting $\alpha(v)$ to the minimum possible value $\forall v \in V$,

$$\textsc{Min-Seed-Cnt}(G, \alpha', 1, r) \geq \textsc{Min-Seed-Cnt}(G, \alpha, 1, r)$$

where α' is any threshold function from $V \to \mathbb{N}$. From Lemma 1. □

Corollary 2 *For every graph G and every valid threshold function α,* $\textsc{Min-Seed-Cnt}(G, \alpha, 1)$ *is bounded below by the* Dominating Set Number *of G.*

Theorem 5 *If there is a $C > 0$ such that a polynomial time algorithm can approximate* $\textsc{Min-Seed-Cnt}(G, \alpha, 1)$ *within* $(1 - C) \ln |V|$, *then* $NP \subseteq TIME(n^{O(\log \log |V|)})$.

Proof: In Theorem 4 we have shown how to reduce every instance of the DOMINATING SET to an instance of $\textsc{Min-Seed-D}(G, \alpha, 1)$ preserving solution size exactly. We know that every instance of the SET COVER problem can be reduced to the DOMINATING SET problem. It is easy to see this is an L-reduction. Feige proved the threshold of $\ln(n)$ approximation for SET COVER [8]. Therefore for dominating set, there is a $\ln(n)$ threshold of approximation (dominating set is equivalent to set cover in terms of approximation ratio) [16, 14]. Hence the theorem. □

4 Approximation Algorithms

4.1 Approximation Algorithm for Min-Seed-Cnt$(G, \alpha, 1)$

We now give an $(H(\alpha_M) + H(\Delta))$ approximation algorithm for $\textsc{Min-Seed-Cnt}(G, \alpha, 1)$.

The approximation algorithm for $\textsc{Min-Seed-Cnt}(G, \alpha, 1)$ is given in Algorithm 1. The technique used in this algorithm is greedy. Say C is the set of already convinced vertices. At each iteration we are choosing a vertex, v, which maximizes $|N[v] \cap (V \setminus C)|$.

Lemma 3 *Algorithm 1 runs in polynomial time.*

Proof: The *while* loop at line number 4 of *Algorithm 1* can execute maximum $|V| - 1$ times. For each $v \in V$ step 6 can take linear time. In worst case *Algorithm 1* takes $\mathcal{O}(n^3)$ time. □

Let the while loop of Algorithm 1 execute s times and let the vertices chosen as the seed be $v_1, v_2, ..., v_s$. Therefore the size of the seed given by Algorithm 1

Algorithm 1 : Algorithm to compute $\textsc{Seed}(G, \alpha, 1)$.

Require: A graph $G = (V, E)$, a function $\alpha : V \to \mathbb{N}$.
$S \leftarrow \emptyset$ {S is the seed}
$C \leftarrow \emptyset$ {C is the set of vertices that are convinced. At the end of the algorithm C must be equal to V}
$i \leftarrow 0$
while $C \neq V$ **do**
  $i \leftarrow i + 1$
  Choose a vertex $v \notin S$ which maximizes $|N[v] \cap (V \setminus C)|$
  $S \leftarrow S \cup \{v\}$
  if $v \notin C_1 \cup C_2 \cup ... \cup C_{i-1}$ **then**
    $C_i \leftarrow \{v\} \cup$ set of vertices newly convinced by choosing v
  else
    $C_i \leftarrow$ set of vertices newly convinced by choosing v
  end if
  $C \leftarrow C \cup C_i$
end while
return S {S is the seed}

is s. Now for $i = 1$ to s and $\forall u \in (N[v_i] \cap (V - C_1 \cup C_2 \cup ... \cup C_{i-1}))$ assign cost $\frac{1}{|N[v_i] \cap (V - C_1 \cup C_2 \cup ... \cup C_{i-1})|}$.

For every $v \in V$, at most $\alpha(v)$ values are assigned. Let the values assigned to v be $d_v^1, d_v^2, ..., d_v^{\alpha(v)}$. Define $c'_v = \sum_{1 \leq i \leq \alpha(v)} d_v^i$ and $c_v = \max_{1 \leq i \leq \alpha(v)} d_v^i$. It follows from the definitions that $\sum_{v \in V} c'_v = S$.

Lemma 4 *$d_v^1 \leq d_v^2 \leq ... \leq d_v^{\alpha(v)}$ and $c'_v \leq \alpha(v) c_v$.*

Proof: From Algorithm 1 it is obvious that, for $1 \leq i < s$

$$|N[v_i] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})| \geq |N[v_{i+1}] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_i)|.$$

Therefore, $\forall v \in V$ $d_v^i \leq d_v^{i+1}$, until d_v^{i+1} is defined and $\forall v \in V \setminus S$, $c_v = d_v^{\alpha(v)}$. Now from the definition of c'_v, we have

$$\forall v \in V,\, c'_v = \sum_{1 \leq i \leq \alpha(v)} d_v^i \leq \alpha(v) c_v.$$

□

Let $|S^*|$ be an optimal seed, so that $\textsc{Min-Seed-Cnt}(G, \alpha, 1) = |S^*|$.

Lemma 5 *For all $v \in V$, $c'_v \leq H(\alpha(v))$.*

Proof: From the definition of c'_v, we know that $c'_v = \sum_{1 \leq i \leq \alpha(v)} d_v^i$. Define α_v^i be the remaining threshold value of v after choosing $v_1, v_2, ..., v_{i-1}$. Therefore,

$$\alpha(v) = \alpha_v^1 \geq \alpha_v^2 \geq ... \geq \alpha_v^s.$$

From the definition of α_v^i, we have

$\alpha_v^{i-1} \leq |N[v] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})| \leq |N[v_i] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})|$.

Therefore,

$$\begin{aligned} c'_v &= \sum_{1 \leq i \leq \alpha(v)} d_v^i = \sum_{1 \leq i \leq s} (\alpha_v^{i-1} - \alpha_v^i) \frac{1}{|N[v_i] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})|} \\ &\leq \sum_{1 \leq i \leq s} (\alpha_v^{i-1} - \alpha_v^i) \frac{1}{\alpha_v^{i-1}} \leq H(\alpha(v)). \end{aligned}$$

□

Lemma 6 $\sum_{v \in V} c'_v \leq \sum_{v \in S^*} \sum_{u \in N(v)} c_u + \sum_{v \in S^*} c'_v$.

Proof: We know that S^* is the optimal solution for Min-Seed-Cnt$(G, \alpha, 1)$. We can divide $\sum_{v \in V} c'_v$ as,

$$\sum_{v \in V} c'_v = \sum_{v \in S^*} c'_v + \sum_{v \in V \setminus S^*} c'_v.$$

From Lemma 4, we have $c'_v \leq \alpha(v) c_v$. Therefore,

$$\sum_{v \in V \setminus S^*} c'_v \leq \sum_{v \in V \setminus S^*} \alpha(v) c_v.$$

As S^* is a seed, $\forall v \in V \setminus S^*$, $|N(v) \cap S^*| \geq \alpha(v)$. Therefore,

$$\begin{aligned} \sum_{v \in V \setminus S^*} \alpha(v) c_v &\leq \sum_{v \in V \setminus S^*} |N(v) \cap S^*| c_v \\ &= \sum_{v \in V \setminus S^*} \sum_{1 \leq i \leq |N(v) \cap S^*|} c_v \\ &= \sum_{u \in S^*} \sum_{v \in N(u) \cap V \setminus S^*} c_v \\ &\leq \sum_{u \in S^*} \sum_{v \in N(u)} c_v. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{v \in V} c'_v &= \sum_{v \in S^*} c'_v + \sum_{v \in V \setminus S^*} c'_v \\ &\leq \sum_{v \in S^*} c'_v + \sum_{v \in V \setminus S^*} \alpha(v) c_v \\ &\leq \sum_{v \in S^*} c'_v + \sum_{u \in S^*} \sum_{v \in N(u)} c_v. \end{aligned}$$

□

Lemma 7 $\forall v \in V, \sum_{u \in N(v)} c_u \leq H(\Delta)$.

Proof: Let v be a vertex in V. Define z_v^i to be the number of unconvinced neighbours of v after the i^{th} stage; that is, the number of nodes in $N(v) \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_i)$, where $1 \leq i \leq s$. Therefore,

$$|N(v)| = z_v^0 \geq z_v^1 \geq z_v^2 \geq ... \geq z_v^s.$$

$(z_v^{i-1} - z_v^i)$ is the number of neighbours of v convinced at stage i. From Algorithm 1 and from the definition of c_v, we have

$$\sum_{u \in N(v)} c_u = \sum_{1 \leq i \leq s} (z_v^{i-1} - z_v^i) \frac{1}{|N[v_i] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})|}.$$

From Algorithm 1, we know that

$$z_v^{i-1} \leq |N[v_i] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})|.$$

Therefore,

$$\begin{aligned} \sum_{u \in N(v)} c_u &= \sum_{1 \leq i \leq s} (z_v^{i-1} - z_v^i) \frac{1}{|N[v_i] \cap (V \setminus C_1 \cup C_2 \cup ... \cup C_{i-1})|} \\ &\leq \sum_{1 \leq i \leq s} (z_v^{i-1} - z_v^i) \frac{1}{z_v^{i-1}} \leq H(|N(v)|) \leq H(\Delta). \end{aligned}$$

□

Theorem 6 *Algorithm 1 is $(H(\alpha_M) + H(\Delta))$ approximation algorithm for* MIN-SEED-CNT$(G, \alpha, 1)$.

Proof: From Lemma 6, we have

$$|S| = \sum_{v \in V} c'_v \leq \sum_{v \in S^*} \sum_{u \in N(v)} c_u + \sum_{v \in S^*} c'_v.$$

From Lemma 5 and Lemma 7, we have

$$c'_v \leq H(\alpha_M) \text{ and } \sum_{u \in N(v)} c_u \leq H(\Delta).$$

Therefore,

$$\begin{aligned} |S| &\leq \sum_{v \in S^*} H(\alpha_M) + \sum_{v \in S^*} H(\Delta) \\ &= (H(\alpha_M) + H(\Delta))|S^*|. \end{aligned}$$

□

4.2 Bounded Degree and p-claw-free Graphs

Theorem 7 MIN-SEED-CNT$(G, \alpha, 1)$ *is $APX - Complete$ for bounded degree graphs.*

Proof: From Theorem 6 if the degree of a graph G is bounded by constant then Algorithm 1 gives a constant approximation ratio. This implies that MIN-SEED-CNT$(G, \alpha, 1)$ belongs to APX. For proving $APX - Completeness$ we have to give an $L - reduction$ from known $APX - complete$ problem. We know that DOMINATING SET on bounded degree graphs is $APX - Complete$ [11]. The reduction used in Theorem 4 to show $NP - complete$ also acts as the reduction to show $APX - Completeness$ with a constant $a = 1$ and $b = 1$ (See Definition 6). □

4.2.1 p-claw free graphs

Lemma 8 *Let $G = (V, E)$ be any p-claw free graph and let $\alpha : V \to \mathbb{N}$ be any threshold function. Let $|D^*_\alpha|$ be an optimal seed, so that* MIN-SEED-CNT$(G, \alpha, 1) = |D^*_\alpha|$ *and S be any Independent Set of G. Let α_m be the minimum threshold value of S vertices and $c = \max\{1, \dfrac{\alpha_m}{p-1}\}$ and $c \in \mathbb{R}^+$. Then $|S| \leq \frac{c(p-1)}{\alpha_m}|D^*_\alpha|$.*

Proof: For all $u \in S \setminus D^*_\alpha$, define $c_u = |(D^*_\alpha \setminus S) \cap N(u)|$. Let α_m be the minimum threshold of S vertices then

$$\sum_{u \in S \setminus D^*_\alpha} c_u \geq \alpha_m |S \setminus D^*_\alpha|$$

For all $v \in D^*_\alpha \setminus S$, define $d_v = |(S \setminus D^*_\alpha) \cap N(v)|$. We know that $\forall v \in D^*_\alpha$ there are at most $(p-1)$ independent vertices in its neighbourhood and $d_v \leq p - 1$.

$$\sum_{v \in D^*_\alpha \setminus S} d_v \leq (p-1)|D^*_\alpha \setminus S|$$

Now consider the definitions of c_u and d_v,

$$\sum_{u \in S \setminus D^*_\alpha} c_u = \{(u, v) \in E \text{ such that } u \in S \setminus D^*_\alpha \text{ and } v \in D^*_\alpha \setminus S\} = \sum_{v \in D^*_\alpha \setminus S} d_v.$$

Therefore,

$$\alpha_m |S \setminus D^*_\alpha| \leq \sum_{u \in S \setminus D^*_\alpha} c_u = \sum_{v \in D^*_\alpha \setminus S} d_v \leq (p-1)|D^*_\alpha \setminus S|$$

$$\begin{aligned}
\alpha_m |S \setminus D^*_\alpha| &\leq (p-1)|D^*_\alpha \setminus S| \\
\alpha_m (|S| - |S \cap D^*_\alpha|) &\leq (p-1)(|D^*_\alpha| - |D^*_\alpha \cap S|) \\
\alpha_m |S| - \alpha_m |S \cap D^*_\alpha| &\leq (p-1)|D^*_\alpha| - (p-1)|D^*_\alpha \cap S| \\
\alpha_m |S| &\leq c(p-1)|D^*_\alpha| \\
|S| &\leq \frac{c(p-1)}{\alpha_m}|D^*_\alpha|.
\end{aligned}$$

□

Algorithm 2 : Algorithm to compute $\textsc{Seed}(G, \alpha, 1)$ in p-claw-free graphs.

Require: A graph $G = (V, E)$, a function $\alpha : V \to \mathbb{N}$.

$D \leftarrow \emptyset$ {D is the seed}
$C \leftarrow \emptyset$ {C is the set of vertices that are convinced.}
$i \leftarrow 0$
while $C \neq V$ **do**
    $i \leftarrow i + 1$
    Choose a Maximal Independent Set S_i from $V \setminus C$
    $C \leftarrow C \cup S_i \cup \{$ set of vertices newly convinced by choosing $S_i\}$
    $D \leftarrow D \cup S_i$
end while
return D {D is the seed}

Lemma 9 *Algorithm 2 is a* $\frac{c\alpha_M(p-1)}{\alpha_m}$ *approximation algorithm for* $\textsc{Min-Seed-Cnt}(G, \alpha, 1)$ *on p-claw free graphs, where* α_M *is the maximum threshold in* G, α_m *is the minimum threshold in* G *and* $c = \max\{1, \dfrac{\alpha_m}{p-1}\}$.

Proof: Let D be the solution given by Algorithm 2 and let D^*_α be the optimal solution for $\textsc{Min-Seed}(G, \alpha, 1)$. Let the while loop of Algorithm 2 execute k times and let the Maximal Independent Sets chosen be $S_1, S_2, ..., S_k$. From Lemma 8, we have

$$|S_i| \leq \tfrac{c(p-1)}{\alpha_m}|D^*_\alpha|, \text{ where } 1 \leq i \leq k.$$

By summation of all Maximal Independent Sets, we have

$$|S_1| + |S_2| + ... + |S_k| \leq \tfrac{kc(p-1)}{\alpha_m}|D^*_\alpha|.$$

From Algorithm 2, we know that $|D| = |S_1| + |S_2| + ... + |S_k|$.
Therefore,

$$|D| \leq \tfrac{kc(p-1)}{\alpha_m}|D^*_\alpha|.$$

At each stage, for every unconvinced vertex u, either u is included in the *seed*, or some unconvinced neighbour of u is included in the maximal independent set. Thus as long as u is unconvinced, at least one neighbour is newly convinced in each stage. Thus every vertex u is convinced in at most $\alpha(u) \leq \alpha_M$ stages. Thus the maximum possible value for k is α_M.
Therefore,

$$|D| \leq \tfrac{c\alpha_M(p-1)}{\alpha_m}|D^*_\alpha|.$$

□

4.3 Spreading Messages with r Radius of Coverage

Theorem 8 *Approximation upper bound for* Min-Seed-Cnt$(G, \alpha, 1, r)$ *is* $(H(n) + H(\alpha_M))$.

Proof: The proof for this is similar to the proof of Theorem 6. The algorithm for this problem will be algorithm 1 by replacing $N(v)$ with $N^r(v)$ and $N[v]$ with $N^r[v]$. The analysis for the approximation is also similar. In the analysis by replacing $N(v)$ with $N^r(v)$ and $N[v]$ with $N^r[v]$ we will get the approximation ratio $(H(n) + H(\alpha_M))$. □

5 Conclusion and Open Problems

In this paper we considered several variants of the Spreading Messages problem and provided complexity results and approximation algorithms for the same.

No approximation algorithm exists for the unbounded spreading messages problem. We have an idea of reducing the instance of Min-Seed-Cnt(G, α, ∞) to Min-Seed-Cnt$(G', \alpha, 1)$. The reduction algorithm is specified below:

Algorithm 3 : Algorithm to reduce Min-Seed-Cnt(G, α, ∞) to Min-Seed-Cnt$(G', \alpha, 1)$.

Require: A graph $G = (V, E)$, a function $\alpha : V \to \mathbb{N}$.

```
Choose a vertex v ∈ V such that ∀u ∈ N(v), α(u) < deg(u).
while Such a v exists do
   Update degrees of N(v) vertices
   V ← V − {v}
   Choose a vertex v ∈ V such that ∀u ∈ N(v), α(u) < deg(u).
end while
```

Let the resultant graph after execution of *Algorithm 3* be $G' = (V', E')$.

Theorem 9 *The feasible solution for* Seed$(G', \alpha, 1)$ *is also a feasible solution for* Seed(G, α, ∞).

Proof: Let all the vertices of the graph G' be convinced. Now consider the vertices that are removed in *Algorithm 3* in reverse order. Remember that the threshold of V' vertices is same as V vertices. So if we add the vertices that are deleted from G in reverse order, the threshold values of V' vertices are not going to change. So, if we add vertices one by one in reverse order they will also get convinced. □

We proved that Min-Seed-D(G, α, ∞, s) is NP-Complete on *Bipartite Graphs*. So, it is also interesting to look at this problem on *Bipartite Permutation Graphs*.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

References

[1] E. Ackerman, O. Ben-Zwi, and G. Wolfovitz. Note: Combinatorial model and bounds for target set selection. *Theor. Comput. Sci.*, 411:4017–4022, October 2010.

[2] P. Alimonti and V. Kann. Hardness of approximating problems on cubic graphs. In *CIAC '97: Proceedings of the Third Italian Conference on Algorithms and Complexity*, pages 288–298, London, UK, 1997. Springer-Verlag.

[3] C.-L. Chang and Y.-D. Lyuu. On irreversible dynamic monopolies in general graphs. *CoRR*, abs/0904.2306, 2009.

[4] C.-L. Chang and Y.-D. Lyuu. Spreading messages. *Theor. Comput. Sci.*, 410(27-29):2714–2724, 2009.

[5] C.-L. Chang and Y.-D. Lyuu. Spreading of messages in random graphs. In R. Downey and P. Manyem, editors, *Fifteenth Computing: The Australasian Theory Symposium (CATS 2009)*, volume 94 of *CRPIT*, pages 3–7, Wellington, New Zealand, 2009. ACS.

[6] N. Chen. On the approximability of influence in social networks. *SIAM J. Discret. Math.*, 23:1400–1415, September 2009.

[7] P. A. Dreyer, Jr. and F. S. Roberts. Irreversible k-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discrete Appl. Math.*, 157(7):1615–1627, 2009.

[8] U. Feige. A threshold of ln for approximating set cover. *J. ACM*, 45(4):634–652, 1998.

[9] P. Flocchini, F. Geurts, and N. Santoro. Optimal irreversible dynamos in chordal rings. *Discrete Applied Mathematics*, 113(1):23–42, 2001.

[10] P. Flocchini, E. Lodi, F. Luccio, and N. Santoro. Irreversible dynamos in tori. In *Euro-Par*, pages 554–562, 1998.

[11] R. Klasing and C. Laforest. Hardness results and approximation algorithms of k-tuple domination in graphs. *Inf. Process. Lett.*, 89(2):75–83, 2004.

[12] C.-S. Liao and G. J. Chang. k-tuple domination in graphs. *Inf. Process. Lett.*, 87(1):45–50, 2003.

[13] F. Luccio, L. Pagli, and H. Sanossian. Irreversible dynamos in butterflies. In C. Gavoille, J.-C. Bermond, and A. Raspaud, editors, *SIROCCO*, pages 204–218. Carleton Scientific, 1999.

[14] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *J. ACM*, 41(5):960–981, 1994.

[15] C. M. Papadimitriou. *Computational complexity.* Addison-Wesley, Reading, Massachusetts, 1994.

[16] A. Paz and S. Moran. Non deterministic polynomial optimization problems and their approximations. *Theor. Comput. Sci.*, 15:251–277, 1981.

[17] D. Peleg. Size bounds for dynamic monopolies. *Discrete Applied Mathematics*, 86(2-3):263–273, 1998.

[18] T. V. T. Reddy, D. S. Krishna, and C. P. Rangan. Variants of spreading messages. In *WALCOM*, volume 5942 of *Lecture Notes in Computer Science*, pages 240–251. Springer, 2010.

[19] D. Watts. A simple model of global cascades on random networks. *P Natl Acad Sci USA*, 99(9):5766–5771, 2002.