

Article

Latency-Bounded Target Set Selection in Signed Networks

Miriam Di Ianni ^{1,*} and Giovanna Varricchio ² ¹ Dipartimento di Matematica, Università degli Studi di Roma “Tor Vergata”, 00133 Roma, RM, Italy² Gran Sasso Science Institute, 67100 L’Aquila, Italy; giovanna.varricchio@gssi.it

* Correspondence: miriam.di.ianni@uniroma2.it

Received: 22 December 2019; Accepted: 27 January 2020; Published: 29 January 2020



Abstract: It is well-documented that social networks play a considerable role in information spreading. The dynamic processes governing the diffusion of information have been studied in many fields, including epidemiology, sociology, economics, and computer science. A widely studied problem in the area of viral marketing is the target set selection: in order to market a new product, hoping it will be adopted by a large fraction of individuals in the network, which set of individuals should we “target” (for instance, by offering them free samples of the product)? In this paper, we introduce a diffusion model in which some of the neighbors of a node have a negative influence on that node, namely, they induce the node to reject the feature that is supposed to be spread. We study the target set selection problem within this model, first proving a strong inapproximability result holding also when the diffusion process is required to reach all the nodes in a couple of rounds. Then, we consider a set of restrictions under which the problem is approximable to some extent.

Keywords: social networks; network analysis; diffusion processes; approximation

1. Introduction

The study of diffusion processes is a widely investigated topic in the complex networks setting. It aims at analyzing how local interactions among nearby nodes may lead to the diffusion, possibly over the whole network, of some feature (opinion/information/disease) starting from a limited set of initiator (or seed) nodes. The problem known as Influence Maximization has been introduced by Domingos and Richardson [1,2] for the design of viral marketing campaigns on social networks by using a small (let us say, given) number of initiators (for example by providing a limited set of people free samples of a product to be merchandised), starting from which a cascade is triggered maximizing the overall spreading of information (see also [3–9]). Dually, the target set selection problem aims at finding out a minimum size set of initiators able to spread the feature over the whole network [3,10–17] or over a large (let us say, given) portion of the network; such a set of initiators is called a target set.

Diffusion processes are usually modeled by graphs in which each vertex may be in one out of two states: informed or unaware. Initially, only nodes in a given set I_0 are informed, while all the others are unaware; while the diffusion process goes on, nodes may change their states according to the diffusion rule of the process. A diffusion rule may be reversible, if changes from informed to unaware are allowed, or irreversible if a node can no longer change its state once it becomes informed. In a discrete diffusion model, nodes (eventually) change their state at discrete time steps. In the rest of this paper, we shall only deal with discrete irreversible diffusion processes.

Several rules to govern the transition from unaware to informed have been proposed, and many papers have studied the dynamics of these systems mostly in stochastic scenarios (see [18,19] and references quoted therein). In the Independent Cascade Model [6,20] an edge-weighted graph is assumed and, as soon as a node u becomes informed, each currently unaware neighbor v of u changes

its state to informed with probability equal to the ratio between the weight of arc (u, v) and the sum of weights of all arcs incident on u . A deterministic counterpart of this model is that of a *Majority Process* [21–23] in which, at each time step, a vertex changes its state if and only if at least half of its neighbors are in the opposite state.

The Linear Threshold Model [5,6,24–26] is based on the association of a threshold $\theta(v)$ to each node v and of a weight $w(u, v)$ to each arc (u, v) : node u becomes informed if and only if $\sum_{v \in N(u) \cap I} w(u, v) \geq \theta(u)$, where $N(u)$ is the set of neighbors of u and I is the set of the already informed nodes. In other words, the Linear Threshold Model assumes that any unaware node becomes informed if the weight of its informed neighbors is above a certain threshold (i.e., the node is subject to a large enough amount of “social pressure”). In [25,27] a special case of the Linear Threshold Model is considered, in which all arc weights are 1 and all nodes have the same threshold θ ; in this case, the threshold value just stands for the number of neighbors that have to be informed in order to induce a node to become informed. In that paper, the authors prove that deciding if a graph has a target set of size d is an NP-complete problem for $\theta \geq 3$. Chen [13] then strengthened such result by showing the NP-hardness of the problem even for $\theta = 2$ and proved that, in the same hypothesis, no polynomial-time algorithm exists with approximation factor better than $O(2 \log^{1-\epsilon} |V|)$ also when all nodes have constant degrees, unless $\text{NP} \subseteq \text{DTIME}(n^{\text{polylog}n})$; in the same paper a polynomial-time algorithm to find an optimal solution has been provided when the underlying graph is a tree. Other interesting cases in which the problems become tractable are studied in [16]. In [3], Ben-Zwi et al. proved that the problem can be solved in time $O(M^w |V|)$ where M is the maximum threshold value and w is the treewidth of the graph: since M cannot exceed the maximum node degree, this proves that the problem is fixed-parameter tractable if parametrized with respect to both treewidth and the maximum node degree of the graph.

The aforementioned papers did not consider the issue of the number of time steps necessary for the feature diffusion. However, this is a relevant point in many settings: in viral marketing, for instance, it is quite important to spread information quickly. Indeed, research in Behavioural Economics shows that humans make decisions mostly on the basis of very recent events [28,29]; furthermore, marketing strategies are more effective when the time needed to reach all the target individuals is short. Hence, it is a worthwhile goal to study diffusion processes while taking into account the number of time steps within which it is required to have all the nodes informed and, actually, this topic has already received some attention. Indeed, as pointed out by the authors, the inapproximability result proved in [13] for general graphs still holds when the diffusion process is required to end in a bounded number of time steps (16 time steps, in fact). In [24] a polynomial-time algorithm spreading a feature over a set of required size within a constant number of time steps in graphs of bounded clique-width has been proposed. Following [24], in this paper, diffusion processes with an available fixed number of time steps within which informing all the nodes will be called latency-bounded diffusion processes.

All diffusion rules considered so far assume that nodes are always positively influenced by their neighbors, that is, any node always has increased its support in favor of the feature being spread throughout the network whenever one of its neighbors becomes informed. However, in most network settings also negative link effects are to be considered.

One way to take into account negative feedbacks is that pursued in [30–33] where it is assumed that individuals may also develop a negative opinion about the feature to be spread and, in this case, they may negatively influence their neighbors. In [30] a generalization of the Independent Cascade model is proposed in which each initiator turns positive (experiencing good quality of the feature) with probability q and turns negative (disliking the feature) with probability $1 - q$, and then, at each time step, a node having become positively informed in the previous step tries to inform each of its unaware neighbors: if successful (with success probability p) the neighbor becomes positively informed with probability q and negatively informed with probability $1 - q$; meanwhile, a negatively informed node in the previous step tries to negatively inform its unaware neighbors, and if successful the neighbors become negatively informed. In [33], instead, a generalization of the Linear Threshold

model is proposed in which each node (which can be inactive, positively active or negatively active) v has a belief score $r(v)$ (measuring the amount of trust that the node places on its own judgment) and a threshold $\theta(v)$, and each arc has a weight $w(u, v)$; a node is activated the first time the sum of weights from its active neighbors plus its own belief score exceeds its threshold and, after being activated, the node decides on its attitude (positive or negative) based on the ratio between the positive influence (sum of weights from positively activated friends) and negative influence (sum of weights from negatively activated friends).

A different approach to negative feedbacks is tackled in [34,35] taking into account the possibility that some relations between individuals are ruled by, for instance, antagonism and distrusting (see [36] and references quoted therein). Actually, in all papers cited so far, it is assumed that nodes always trust their neighbors and, hence, they get support in favor to accept the feature by their neighbors having a positive opinion about it and they get support in favor to discard the feature by their neighbors having a negative opinion about it. Conversely, it is now assumed that node opinions about the feature to be spread are always positive but that receiving positive feedback from an untrusted/antagonist neighbor results in increasing the support to discard the feature. Ahmed et al. [34] presented a new threshold model to incorporate positive and negative influences among individuals in which the effect of negative influence is the tendency of individuals in not adopting a product adopted by their opponent: each arc (u, v) has a weight $p(u, v) \in [0, 1]$, each node v has a couple of threshold values, $\theta^+(v)$ and $\theta^-(v)$ and, at any time step, if S^+ and S^- are, respectively, the sets of the trusted and untrusted active neighbors of v at that time step, v becomes active if $1 - \prod_{u \in S^+} (1 - p(u, v)) > \theta^+(v)$ and $1 - \prod_{u \in S^-} (1 - p(u, v)) < \theta^-(v)$. In [35], echoing the principles that “the friend of my enemy is my enemy” and “the enemy of my enemy is my friend”, a reversible model is presented in which positive relationships carry the influence in a positive manner, i.e., you would more likely trust and adopt your friends’ opinions, while negative relationships often carry influence in a reverse direction (if your foe chooses one opinion or votes for one candidate, you would more likely be influenced to do the opposite): at each step, every node randomly picks one of its neighbors, and if the arc from this neighbor is positive the node adopts the neighbor’s decision (to accept the feature or not), while if the edge is negative the node adopts the opposite of the neighbor’s decision.

In [37,38] the two approaches (positive/negative opinions and positive/negative relationships) are jointly considered. In [37], an extension of the independent cascade model is proposed in which each node u has an opinion $o_u \in [-1, 1]$, each link (u, v) has a probability of interaction $\varphi_{u,v} \in [0, 1]$ and a probability of activation $p(u, v)$: as soon as u becomes active, it gets a unique chance of turning v active with probability $p(u, v)$; if v is activated by u , then the probability that v gets informed with the same orientation of u is $\varphi_{u,v}$, and the probability that v gets informed with the opposite orientation is $1 - \varphi_{u,v}$. In [38] another extension of the independent cascade model is introduced in which nodes may be in one of the three states: positively active, negatively active or inactive; when node v becomes somehow active, it tries to activate each one of its inactive neighbor nodes with its same opinion (positive if it is positively active or negative if it is negatively active) in the case its relation with that neighbor is positive, with opposite opinion in the case its relation with that neighbor is negative. The interest of most of the cited papers is in the Influence Maximization problem with the goal, in case positive/negative opinions are considered, of maximizing the spread of positive opinions. We also notice that most of the proposed models taking into account negative feedback effects are stochastic.

In this paper we follow the same approach to negative feedback effects as in [34,35], introducing a deterministic diffusion model in which nodes’ opinions about the feature to be spread are always positive but the relationships between nodes may be positive or negative. In more detail, we consider a generalization of the Linear Threshold Model in which the set of neighbors of each node is partitioned into two subsets: one subset of trusty neighbors (or friends), containing the neighbors able to positively influence the node, and one subset of unreliable neighbors (or enemies), containing the neighbors able to negatively influence the node. Informally speaking, in our model an unaware node becomes informed at some time step only if at that time step the cumulative influence coming from its informed

trustworthy neighbors is at least as large as the node threshold, and at that time step as well as at any previous time step the cumulative influence coming from its informed unreliable neighbors is less than the node threshold. Notice that, as in [30], positive and negative influences are slightly asymmetric in that negative feedback effects are more powerful than positive ones in influencing a node: indeed, in [30] a negatively informed node always tries to negatively inform its unaware neighbors (while there is a chance that a node gets negatively informed by a positively informed neighbor), and in our model an unaware node will never become informed if at some time step the cumulative influence from its enemies exceeds the threshold (also if in the same time step the cumulative influence from its friends exceeds the threshold too). In fact, several studies in social psychology (e.g., [39–42]) point out that negative impact is usually stronger and much more dominant than positive impact in shaping people’s decisions. Notice also that our rule according to which an unaware node becomes informed is quite similar to that in [34], discussed above (although [34] deals with the Influence Maximization problem). The main difference between the two rules is that, while the rule herein defined directly generalizes that in the Linear Threshold model in its considering the cumulative amount of information reaching an unaware node (i.e., we sum up the weights of informed neighbors), the rule in [34] is somehow related to that in the Independent Cascade model: in fact, in that paper, the function associated to the set S^+ of the informed friends (and similarly as to the informed enemies) is $1 - \prod_{u \in S^+} (1 - p(u, v))$ which denotes the probability of accepting, let’s say, the suggestion coming from at least one informed friend in the hypothesis of independent probabilities of accepting the suggestion from any of them.

Within this setting, we consider a problem that is closely related to that studied in [24], that is, computing a minimum size set able to spread a feature to all the network nodes within a given number k of time steps. We shall call such a set a k -target set.

2. Results and Discussion

After providing the formal definitions in Section 3, in Section 4 we prove that computing both a minimum size 2-target set and a minimum size target set (of unbounded latency) with respect to our diffusion rule is MIN HORN DELETION-complete with respect to a reduction preserving approximability properties of (NPO) optimization problems (we shortly and informally recall that a problem \mathcal{P} is \mathcal{Q} -complete with respect to some reduction α , for some problem \mathcal{Q} , if \mathcal{P} is α -reducible to \mathcal{Q} and \mathcal{P} is α -reducible to \mathcal{Q} . If this happens, for any property π preserved by α , \mathcal{P} satisfies π if and only if \mathcal{Q} satisfies π . Needless to say, \mathcal{P} is \mathcal{Q} -complete if and only if \mathcal{Q} is \mathcal{P} -complete). Hence, the two problems are both contained in the class poly-APX and NP-hard to approximate to within $2^{\log^{1-\epsilon} |V|}$ for any $\epsilon > 0$, also when node thresholds and arc weights are all equal to 1, that is, when the diffusion rule is a direct generalization of the one in [25].

In Section 5, we start the study of the problem in which a minimum size 1-target set has to be computed. Such a problem is trivially SET COVER-HARD also when node thresholds and arc weights are all equal to 1. We start by showing that the same reduction used to upper bound the complexity of the unbounded latency and of the latency 2 cases actually yields a $(\Delta^+ + 1)$ -approximation for the latency 1 diffusion processes, where Δ^+ is the maximum in-degree in the graph of the positive relations. We notice that the $(\Delta^+ + 1)$ -approximation does not depend in any way on the topology of the graph of the negative relations.

The study of the way the complexity of the search for 1-target sets is related to the topology of the graph of the negative relations is started in Section 5.1: we first show the problem is in PO when restricted to instances such that the graph of the negative relations is a generalized chain (that is, the graph obtained by replacing each of its strongly connected components by a node is a directed chain), then we get a non-polynomial time $(\log |V| + 1)$ -approximation algorithm for general graph by reducing to WEIGHTED SET COVER. Finally, we show that such algorithm runs in polynomial time when the graph of the negative relations is a generalized source tree, that is when the graph obtained by replacing its strongly connected components by nodes is a source tree, thus yielding a polynomial-time $(\log |V| + 1)$ -approximation algorithm for generalized source trees. It is worthwhile

to remark that the results in Section 5.1 hold whatever topology of the graph of the positive relations is considered.

We observe that the study of diffusion processes of duration 1 is also of some interest in the graph-theoretic setting. In fact, the search for a minimum 1-target set with respect to the Linear Threshold rule can be seen also as the search for some kind of dominating set. As an example, 1-target sets with respect to the diffusion rule in [25], in which an unaware node becomes informed as soon as θ of its neighbors are informed, are just θ -dominating sets [43]. Similarly, 1-target sets with respect to our unweighted diffusion rule correspond to a sort of *dominating sets with constrained pairs*: given a graph $G = (V, E)$ and a set C of pairs of nodes, a constrained dominating set is $D \subseteq V$ of nodes such that *i*) for any $u \in V - D$ there is $v \in V$ such that $(v, u) \in E$ and *ii*) for any $(u, v) \in C$ either $u \in D$ and $v \in D$ or $u \notin D$ and $v \notin D$.

3. Model and Problems Definitions

A signed graph $G = (V, A^+ \cup A^-)$ is the overlapping of two directed graphs, $G^+ = (V, A^+)$ and $G^- = (V, A^-)$, over the same set of nodes. For an example of the definitions to follow, refer to Figure 1.

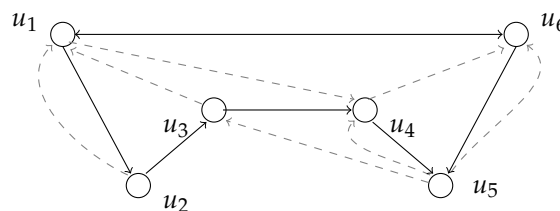


Figure 1. A signed graph $G = (V, A^+ \cup A^-)$. Arcs in A^+ are black, arcs in A^- are dashed. In this example, $N_{in}^+(u_3) = \{u_2\}$, $N_{in}^-(u_3) = \{u_5\}$, $N_{out}^+(u_3) = \{u_4\}$, $N_{out}^-(u_3) = \{u_1\}$, $\pi^+(u_3) = \{u_1, u_2, u_6\}$ and $\pi^-(u_3) = \{u_5\}$. Again, $d^+(u_1, u_5) = 2$, $d^-(u_5, u_6) = 1$, and, by setting $V' = \{u_1, u_5\}$, $d^+(V', u_3) = 2$ and $d^-(V', u_3) = 1$. Finally, $D^+ = 4$ (a maximum length positive shortest path being $\langle u_6, u_1, u_2, u_3, u_4 \rangle$) and $D^- = 3$ (a maximum length negative shortest path being $\langle u_2, u_1, u_4, u_6 \rangle$).

For any $u \in V$, we denote as $N_{in}^+(u)$ and $N_{in}^-(u)$, respectively, the sets of *in-friends* and of *in-enemies* of u , that is, the in-neighbors of u in G^+ and in G^- :

$$N_{in}^+(u) = \{v \in V : (v, u) \in A^+\} \quad \text{and} \quad N_{in}^-(u) = \{v \in V : (v, u) \in A^-\}.$$

Similarly, for each $u \in V$, we define the sets of out-friends and out-enemies of u as

$$N_{out}^+(u) = \{v \in V : (u, v) \in A^+\} \quad \text{and} \quad N_{out}^-(u) = \{v \in V : (u, v) \in A^-\}.$$

A path in G^+ between any pair of nodes in V is called a positive path and D^+ denotes the maximum length of a shortest positive path. For $u \in V$, we denote as $\pi^+(u)$ the set of nodes such that, for any $v \in \pi^+(u)$, G contains a positive path from v to u and, for $v \in \pi^+(u)$, as $d^+(v, u)$ the length of a shortest positive path in G from v to u . Finally, for $u \in V$ and $V' \subseteq V$,

$$d^+(V', u) = \begin{cases} \min\{d^+(v, u) : v \in V' \cap \pi^+(u)\} & \text{if } V' \cap \pi^+(u) \neq \emptyset, \\ D^+ + 1 & \text{otherwise.} \end{cases}$$

Negative paths, D^- , $\pi^-(u)$, $d^-(u, v)$ and $d^-(V', u)$ are defined similarly.

Let $\theta : V \rightarrow \mathbb{N}^+$ and $w : A^+ \cup A^- \rightarrow \mathbb{N}^+$ be, respectively, a node-weight and an arc-weight function. A set $I_0 \subseteq V$ of initially informed nodes (all the remaining nodes in $V - I_0$ being unaware) may eventually start a (θ, w) -diffusion process: at time step 1 any node $u \in V - I_0$ becomes informed if

the cumulative weight of its in-friends is at least $\theta(u)$ while the cumulative weight of its in-enemies is less than $\theta(u)$, that is

$$\sum_{v \in N_{in}^+(u) \cap I_0} w(v, u) \geq \theta(u) \quad \text{and} \quad \sum_{v \in N_{in}^-(u) \cap I_0} w(v, u) < \theta(u);$$

We denote as I_1 the set of nodes becoming informed at time step 1.

Conversely, if the cumulative weight of u 's in-enemies is at least $\theta(u)$, then the negative influence that u receives from $N_{in}^-(u)$ will induce it to remain in the unaware state forever, that is, at any successive time step u will never become informed. We say u becomes refractory. Needless to say, if neither the cumulative weight of u 's in-friends nor the cumulative weight of u 's in-enemies reaches $\theta(u)$, u does not change its state.

Inductively, if I_j denotes the set of nodes becoming informed at time step j , an unaware node $u \in V - \cup_{j=0}^{t-1} I_j$ becomes refractory at time step $t > 1$ if

$$\sum_{v \in N_{in}^-(u) \cap (I_0 \cup I_1 \cup \dots \cup I_{t-1})} w(v, u) \geq \theta(u), \tag{1}$$

while it becomes informed if (1) does not hold and

$$\sum_{v \in N_{in}^+(u) \cap (I_0 \cup I_1 \cup \dots \cup I_{t-1})} w(v, u) \geq \theta(u).$$

A set $I_0 \subseteq V$ is a *target set* for $\langle G = (V, A^+ \cup A^-), \theta, w \rangle$ if $\cup_{t \geq 0} I_t = V$. Needless to say, the size of a target set $I_0 \subseteq V$ is $|I_0|$.

The SIGNED TARGET SET SELECTION problem (STSS) asks for computing a minimum size target set for a given signed graph $G = (V, A^+ \cup A^-)$ with node- and arc-weight functions θ and w . The UNWEIGHTED SIGNED TARGET SET SELECTION problem (USTSS) is the restriction of STSS in which all node- and arc-weights equal 1.

Notice that, if $I_t = \emptyset$ for some $t \geq 0$, then $I_s = \emptyset$ for all $s > t$. The *latency* of a diffusion process is, therefore, defined as the integer $t \in \mathbb{N}^+$ such that $I_t \neq \emptyset$ and $I_{t+1} = \emptyset$. The t -TSS (t -USTSS) problem asks for computing minimum size target sets yielding diffusion processes of latency t .

4. Complexity of Approximation of USTSS and 2-USTSS

We shortly recall that an A-reduction (see [44]) is a reduction between a pair of (NPO) optimization problems transforming (in polynomial time) instances of a problem P to instances of a problem Q so that a solution for an instance x_P of P can be recovered in polynomial time from a solution x_Q of the corresponding instance of Q and, for some $\alpha > 0$, a r -approximate solution for x_Q is a αr -approximate solution for x_P . Informally speaking, A-reductions preserve the approximability properties of (NPO) optimization problems. In this section, we shall prove that 2-USTSS and USTSS are both MIN HORN DELETION-complete with respect to A-reductions. Such a result will follow by a couple of A-reductions involving the WEAKLY POSITIVE-MINIMUM ONES problem proved to be MIN HORN DELETION-complete with respect to A-reductions in [44].

In more detail, we first A-reduce the WEAKLY POSITIVE-MINIMUM ONES problem to 2-USTSS in Section 4.1, and this together with the results in [44], and Lemmas 1 and 2, will prove the following theorem.

Theorem 1. For any $\epsilon > 0$, 2-USTSS (and, hence, USTSS) is NP-hard to approximate to within $2^{\log^{1-\epsilon} |V|}$.

Then, in Section 4.2 we A-reduce USTSS to WEAKLY POSITIVE-MINIMUM ONES in order to get the MIN HORN DELETION-completeness of both USTSS and 2-USTSS. In turn, this implies that 2-USTSS and USTSS are in poly-APX.

Let $X = \{x_1, \dots, x_n\}$ be a set of boolean variables; a set of clauses (i.e., disjunction of literals) $f = \{c_1, \dots, c_m\}$ on X such that, for each $c_j \in f$, c_j contains at most one negative variable is said *weakly positive*. If f is a set of weakly positive clauses, a clause in f will be called positive if all literals it contains are positive variables, and it will be called negative if it contains one negated variable.

An instance of WEAKLY POSITIVE-MINIMUM ONES is a pair $\langle X, f \rangle$, where $X = \{x_1, \dots, x_n\}$ is a set of boolean variables and $f = \{c_1, \dots, c_m\}$ is a weakly positive set of clauses; the WEAKLY POSITIVE-MINIMUM ONES problem asks for computing a minimum size subset of X to be set to true in order to satisfy all clauses in f .

In what follows, with a slight abuse of notation, we shall sometimes write $x_i \in c_j$ ($\neg x_i \in c_j$) as a shorthand to mean that x_i ($\neg x_i$) is a literal of clause c_j (so that c_j is considered as a set of literals, instead of their disjunction). We shall say that a variable x_i is positively contained in a clause c_j if $x_i \in c_j$, and that x_i is negatively contained in c_j if $\neg x_i \in c_j$.

Let $\langle X, f \rangle$ be an instance of WEAKLY POSITIVE-MINIMUM ONES; without loss of generality, in what follows we shall always assume that at least one clause in f is positive (otherwise all clauses in f would be satisfied by setting to true no variable in X) and that each clause in f contains at least a couple of variables (one of which positive). Indeed, if some clause in f consists of a single positive variable then such variable has to be necessarily set to true in order to satisfy f , and a smaller set of clauses f' can be derived by removing from f all clauses containing that variable and an optimum solution for f is obtained by adding the variable to an optimum solution for f' . Similarly, if some clause in f consists of a single negative variable then such variable has to be necessarily set to false in order to satisfy f , and a smaller set of clauses f' can be derived by removing from f all clauses containing that negative variable and an optimum solution for f is an optimum solution for f' . Hence, f is satisfied by assigning true to all the n variables in X , and f is not satisfied by assigning false to all the n variables in X .

4.1. WEAKLY POSITIVE-MINIMUM ONES \leq_A 2-USTSS

We now derive from an instance $\langle X, f \rangle$ of WEAKLY POSITIVE-MINIMUM ONES a signed network $G = (V, A^+ \cup A^-)$ (see Figure 2):

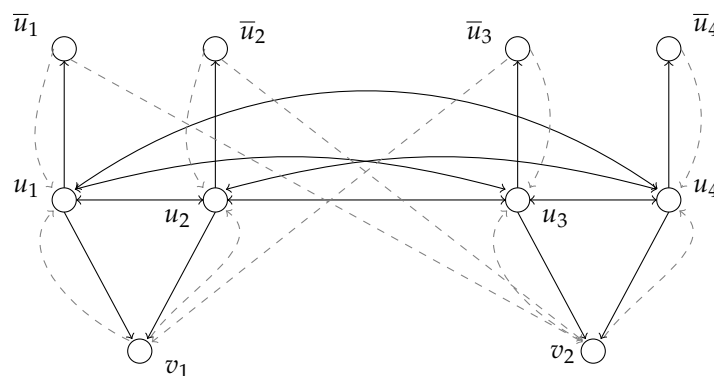


Figure 2. Reducing WEAKLY POSITIVE MINIMUM ONES to USTSS: $f = \{c_1, c_2\}$, with $c_1 = \neg x_3 \vee x_1 \vee x_2$ and $c_2 = x_3 \vee x_4$. Arcs in A^- are dashed.

- we set $V = V_X \cup V_{\bar{X}} \cup V_f$, where $V_X = \{u_i : x_i \in X\}$, $V_{\bar{X}} = \{\bar{u}_i : x_i \in X\}$ and $V_f = \{v_j : c_j \in f\}$;
- for each $x_i \in X$, (u_i, \bar{u}_i) is an arc in A^+ and (\bar{u}_i, u_i) is an arc in A^- ;
- for each $u_i, u_j \in V_X$, (u_i, u_j) is an arc in A^+ and (u_j, u_i) is an arc in A^+ (that is, nodes in V_X form a clique of positive arcs);

- for each clause c_j and for each x_i positively contained in c_j , (u_i, v_j) is an arc in A^+ and (v_j, u_i) is an arc in A^- ;
- if c_j is a negative clause and $\neg x_h$ is its negative variable, then (\bar{u}_h, v_j) is an arc in A^- ;
- if $c_j \in f$ is a positive clause then, for any $x_i \in X - \{x_k : x_k \in c_j\}$, (\bar{u}_i, v_j) is an arc in A^- .

Needless to say, computing G requires polynomial time in $|f|$.

The next lemma proves that it suffices to search in V_X for a target set for G .

Lemma 1. For any target set $I \subseteq V$ for G there exists a target set I_0 for G such that $|I_0| \leq |I|$ and $I_0 \subseteq V_X$.

Proof. Recall that each clause in f contains at least one positive variable.

Let $I \subseteq V$ be a target set for G . Notice that, since all in-friends of any node in V_X are contained in V_X (that is, $N_{in}^+(u) \subseteq V_X$ for any $u \in V_X$), then nodes in V_X can get informed by nodes in V_X only; consequently, it must be $I \cap V_X \neq \emptyset$.

Suppose there exists $i \in \{1, \dots, n\}$ such that $\bar{u}_i \in I$: since $(\bar{u}_i, u_i) \in A^-$, it must be $u_i \in I$ (or u_i could not be informed, contradicting that I is a target set). On the other hand, since $(u_i, \bar{u}_i) \in A^+$ and $N_{in}^-(\bar{u}_i) = \emptyset$, u_i is able to inform \bar{u}_i . Finally, since $N_{out}^+(\bar{u}_i) = \emptyset$, \bar{u}_i is unable to inform any node, this proves that $I - \{\bar{u}_i\}$ is still a target set for G . So, $I_1 = I - V_{\bar{X}}$ is a target set for G .

Similarly, if there exists $v_j \in V_f \cap I_1$, then, since $(v_j, u_i) \in A^-$ for any $x_i \in X$ positively contained in c_j , I_1 is a target set for G only if $u_i \in I_1$. On the other hand, since $(u_i, v_j) \in A^+$, since $N_{in}^-(v_j) \subseteq V_{\bar{X}}$, and since $I_1 \cap V_{\bar{X}} = \emptyset$, then u_i is able to inform v_j . Since $N_{out}^+(v_j) = \emptyset$, v_j is unable to inform any other node, and, hence, $I_1 - \{v_j\}$ is still a target set for G , and so is $I_0 = I_1 - V_f$. \square

The next lemma shows that the construction defined in this subsection is actually an A-reduction, so completing the proof that WEAKLY POSITIVE-MIN ONES \leq_A USTSS.

Lemma 2. For any $k \in \mathbb{N}^+$, with $1 \leq k \leq n$, f is satisfiable by assigning true to k variables in X if and only if there exists a target set for G of size k .

Proof. Suppose f is satisfiable by the truth assignment a assigning true to $k \geq 1$ variables in X . We show that the set $I_0 = \{u_i \in V_X : a(x_i) = \text{true}\}$ is a target set for G .

We start by noticing that all nodes in $V_X - I_0$ become informed at time step 1, and that all nodes in $V_{\bar{X}}$ become informed at time step 1 or 2 so that they cannot negatively influence any node at time step 1. Since a satisfies f , any positive clause c_j contains a (positive) variable x_i such that $a(x_i) = \text{true}$; hence, $u_i \in I_0$ and v_j becomes informed at time step 1. Hence, all nodes in V_f corresponding to positive clauses become informed at time step 1.

Let c_j be a negative clause and let u_h be its negative variable: a satisfies c_j either by assigning false to u_h or by assigning true both to its negative variable and to at least one of its positive variables. In the former case, \bar{u}_h becomes informed at time step 2 and, since all nodes corresponding to positive variables in c_j become informed at time step 0 or 1 (whether they are in I_0 or not), it cannot forbid v_j to get informed at time step 2. In the latter case, both v_j and u_h become informed at time step 1 so that \bar{u}_h cannot forbid v_j to get informed.

Hence, all nodes in G are informed (by time step 2), that is, I_0 is a target set for G , and $|I_0| = k$.

Conversely, let $I_0 \subseteq V$ be a target set for G . By Lemma 1, we can assume $I_0 \subseteq V_X$.

$$a(x_i) = \begin{cases} \text{true} & \text{if } u_i \in I_0, \\ \text{false} & \text{if } u_i \notin I_0. \end{cases}$$

Let c_j be a positive clause in f . If $u_i \notin I_0$ for any $x_i \in c_j$, then v_j could become informed only at time step 2; since $I_0 \subseteq V_X$, then there exists $u_\ell \in I_0 - c_j$ so that \bar{u}_ℓ is informed at time step 1 and, since $(\bar{u}_\ell, v_j) \in A^-$, it would forbid v_j to become informed, contradicting that I_0 is a target set. Hence, for any positive clause c_j in f there exists $x_i \in c_j$ such that $a(x_i) = \text{true}$.

Let c_j be a negative clause in f , let x_h be its negative variable and let x_{j_1}, \dots, x_{j_q} be its positive variables. If $a(x_{j_i}) = \text{false}$ for any $i = 1, \dots, q$, and, hence, $\{u_{j_1}, \dots, u_{j_q}\} \cap I_0 = \emptyset$, then, as already noticed, v_j could become informed only at time step 2. In this case it must be $u_h \notin I_0$: indeed, if $u_h \in I_0$ then \bar{u}_h would get informed at time step 1 and, since $(\bar{u}_h, v_j) \in A^-$, this would forbid v_j to become informed, so contradicting that I_0 is a target set for G . Hence, if $a(x_{j_i}) = \text{false}$ for any $1 \leq i \leq q$, then it must be $a(x_h) = \text{false}$ too.

Summarizing, a satisfies each clause in f , that is, a satisfies f , and a assigns true to $|I_0|$ variables. \square

We finally observe that the latency of any cascade occurring in the signed graph in the reduction here described is at most 2. This proves that, for any $t \geq 2$, t -USTSS is MIN HORN DELETION-hard.

4.2. USTSS \leq_A WEAKLY POSITIVE-MINIMUM ONES

We now derive from a signed graph $G = (V, A^+ \cup A^-)$ a set f_G of weakly positive clauses. Before proceeding, we need a preliminary lemma.

Lemma 3. *A set $I_0 \subseteq V$ is a target set for $G = (V, A^+ \cup A^-)$ if and only if for any $u \in V - I_0$*

- $\pi^+(u) \cap I_0 \neq \emptyset$, and
- for any $v \in N_{in}^-(u)$, it holds that $d^+(I_0, u) \leq d^+(I_0, v)$.

Proof. We start by noticing that, since node thresholds and arc weights are all equal to 1, the diameter D^+ of G^+ upper bounds the latency of any diffusion process in G : indeed, for any $u \in V$, if any of the u 's neighbors is informed at time step t , either u becomes informed at time step $t + 1$ or it will never become informed. As a consequence, for any target set I_0 and for any $u \in V$, u may become informed at time step $t \leq D^+$ only if $d^+(I_0, u) = t$.

Suppose $I_0 \subseteq V$ is a target set for G ; hence, $V - I_0$ can be partitioned into at most D^+ subsets I_1, \dots, I_{D^+} such that all individual-nodes in I_t become informed at time step t , for any $1 \leq t \leq D^+$. This implies that, for any $t = 1, \dots, D^+$ and for any $u \in I_t$, $d^+(I_0, u) = t$ and, hence, $\pi^+(u) \cap I_0 \neq \emptyset$. Let $u \in I_t$, for some $t = 1, \dots, D^+$: if there exists $v \in N_{in}^-(u) \cap I_s$ for some $s \leq t - 1$, v would forbid u to get influenced at time step t . Hence, for any $v \in N_{in}^-(u)$, it must hold that $v \in \cup_{s \geq t} I_s$, that is, $d^+(I_0, u) \leq d^+(I_0, v)$. This proves the first part of the assertion.

Conversely, suppose that, for any $u \in V - I_0$, $\pi^+(u) \cap I_0 \neq \emptyset$ and, for any $v \in N_{in}^-(u)$, it holds that $d^+(I_0, u) \leq d^+(I_0, v)$. For any $1 \leq t \leq D^+$, denote as V_t the set of nodes $u \in V - I_0$ such that $d^+(I_0, u) = t$. Since for any node $u \in V_1$, $N_{in}^-(u) \cap I_0 = \emptyset$ by hypothesis, all nodes in V_1 are influenced at time step 1. Inductively, it can then be proved that, for any $t \leq D^+$, all nodes in V_t , are influenced at time step t . This proves that I_0 is a target set for G . \square

To model $G = (V, A^+ \cup A^-)$ by a weakly positive set of clauses f_G , we start by defining the set $X = \{x_u : u \in V\}$ of boolean variables, each of which witnesses whether the corresponding node is initially informed or not. Then, we describe the set of clauses defining f_G , which is partitioned into a set P of positive clauses and a set Q of negative clauses: that is, $f_G = P \cup Q$.

Let u be any node and let $\pi^+(u)$ be the set $\{v_1, \dots, v_j\}$. P contains a single clause c_u associated to u which models the requirement that u is initially informed or it can become informed in a further step:

$$c_u = x_u \vee x_{v_1} \vee x_{v_2} \vee \dots \vee x_{v_j}.$$

Similarly, for any $u \in V$, Q contains a set Q_u of clauses associated to u , modeling the requirement that u is initially informed or it can become informed earlier than any node in $N_{in}^-(u)$. For each $v \in N_{in}^-(u)$, Q_u contains the clause

$$(\neg x_v \vee x_u) \tag{2}$$

and, for each $1 \leq t \leq D^+$ and for each $z \in \pi^+(v)$ such that $d^+(z, v) = t$, Q_u contains the clause

$$(\neg x_z \vee x_u \vee x_{w_1} \vee x_{w_2} \vee \dots \vee x_{w_k}),$$

where $\{w_1, w_2, \dots, w_k\} = \{w \in V : d^+(w, u) \leq t\}$.

Summarizing, $f_G = P \cup Q$, with $P = \{c_u : u \in V\}$ and $Q = \cup_{u \in V} Q_u$. Needless to say, f_G is computable in polynomial time from G . Furthermore, there is a one-to-one correspondence among subsets of V and truth assignments for X : for any $V' \subseteq V$, we denote as $a_{V'}$ the truth assignment corresponding to V' , that is,

$$a_{V'}(x_u) = \begin{cases} \text{true} & \text{if } u \in V', \\ \text{false} & \text{if } u \in V - V'. \end{cases}$$

The next lemma proves that the just described construction is actually an A-reduction from USTSS to WEAKLY POSITIVE-MINIMUM ONES.

Lemma 4. *Let $G = (V, A^+ \cup A^-)$ be a signed graph. $I_0 \subseteq V$ is a target set for G if and only if the truth assignment a_{I_0} satisfies f_G .*

Proof. Let $I_0 \subseteq V$ be a target set for G .

For any $u \in I_0$, since $x_u \in c_u$ and, for any clause $c \in Q_u$, $x_u \in c$, then a_{I_0} satisfies c_u and every clause in Q_u .

For any $u \in V - I_0$, since I_0 is a target set for G , then, by Lemma 3, $\pi^+(u) \cap I_0 \neq \emptyset$: let $z_u \in \pi^+(u) \cap I_0$ be such that $d^+(z_u, u) = \min\{d^+(w, u) : w \in \pi^+(u) \cap I_0\}$. By construction, $x_{z_u} \in c_u$, and, hence, c_u is satisfied by a_{I_0} . Furthermore, for any $v \in N_{in}^-(u)$, let $y_v \in \pi^+(v) \cap I_0$ be such that $d^+(y_v, v) = \min\{d^+(w, v) : w \in \pi^+(v) \cap I_0\}$: by Lemma 3, for any $v \in N_{in}^-(u)$, $d^+(z_u, u) \leq d_V(y_v, v)$ and, hence, for any $v \in N_{in}^-(u)$, and for any $w \in \pi^+(v) \cap I_0$, x_{z_u} is contained in every clause in Q_u containing $\neg x_w$. As a consequence, every clause $c \in Q_u$ containing $\neg x_w$ as its negative variable such that $a_{I_0}(x_w) = \text{true}$, also contains x_{z_u} ; since $z_u \in I_0$, $a(x_{z_u}) = \text{true}$ and, hence, c is satisfied by a_{I_0} . This proves that every clause $c \in Q_u$ containing $\neg x_w$ as its negative variable such that $a(x_w) = \text{false}$ is satisfied by a .

In conclusion, for any $u \in V$, a_{I_0} satisfies c_u and all clauses in Q_u .

Conversely, suppose a_{I_0} satisfies f_G . For any $u \in V - I_0$, since a_{I_0} satisfies c_u , there must exist at least one node $z \in \pi^+(u)$ such that $a(x_z) = \text{true}$ and, hence, $z \in I_0$. This proves that $\pi^+(u) \cap I_0 \neq \emptyset$. Furthermore, for any $u \in V - I_0$ and $v \in N_{in}^-(u)$, and for any $y \in \pi^+(v) \cap I_0$ (that is, $a_{I_0}(x_y) = \text{true}$ and v may become informed by a path from y), any clause in Q_u containing $\neg x_y$ has also to contain a positive variable x_z such that $a_{I_0}(x_z) = \text{true}$. By construction of the clauses in Q_u , this implies that $d^+(z, u) \leq d^+(y, v)$, that is, $d^+(I_0, u) \leq d^+(I_0, v)$.

By Lemma 3, this proves that I_0 is a target set for G . \square

5. The 1-USTSS Problem

The 1-USTSS problem is trivially SET COVER-hard: actually, a dominating set of a graph $G = (V, E)$ is also a 1-target set for the signed graph $G' = (V, A^+ \cup A^-)$ where $A^+ = \{(u, v), (v, u) : \{u, v\} \in E\}$ and $A^- = \emptyset$. Hence, 1-USTSS is non-approximable within $c \log |V|$ for some $c > 0$.

In the remainder of this section, we study the approximability of 1-USTSS.

We start by noticing that if a diffusion process of latency 1 is required to occur, then the A-reduction described in Section 4.2 yields a weakly positive set of clauses such that all its negative clauses have size 2 (that is, they contain exactly one positive literal and one negative literal). In fact, it directly follows from Lemma 3 that $I_0 \subseteq V$ is 1-target set for a signed graph $G = (V, A^+ \cup A^-)$ if and only if for any $u \in V - I_0$

1. $N_{in}^+(u) \cap I_0 \neq \emptyset$, and

2. for any $v \in N_{in}^-(u)$, it holds that $v \in V - I_0$.

As a consequence, for any $u \in V$ the set of the negative clauses associated to u (as defined in Section 4.2) is

$$Q_u = \{(-x_v \vee x_u) : v \in N_{in}^-(u)\}. \tag{3}$$

This proves that the A-reduction shown in Section 4.2 works also as an A-reduction from 1-USTSS to WEAKLY POSITIVE MINIMUM ONES such that the set of clauses corresponding to a signed graph, besides being weakly positive, has all its negative clauses of size 2. In [44] it is shown that the WEAKLY POSITIVE MINIMUM ONES problem is B -approximable when restricted to such instances, where B is the maximum size of a positive clause in the set.

The size of the positive clauses in the set f_G described in Section 4.2 is very large: for $u \in V$, $|c_u| = 1 + |\pi^+(u)|$, since c_u had to state that u is in I_0 or there exists some node in I_0 connected to u by a path (of any length). However, when diffusion processes of latency 1 are to be described, f_G can be slightly modified in order to decrease the size of its positive clauses: more precisely, condition 1 above implies that, for each $u \in V$, $c_u = x_u \vee x_{v_1} \vee \dots \vee x_{v_j}$ where $\{v_1, \dots, v_j\} = N_{in}^+(u)$, that is, $|c_u| = |N_{in}^+(u)| + 1$.

Hence, by denoting as Δ^+ the maximum in-degree of a node in G^+ , that is, $\Delta^+ = \max\{|N_{in}^+(u)| : u \in V\}$, the next theorem follows from Lemma 4, from definition (3) and from the already cited result in [44].

Theorem 2. *The 1-USTSS problem is $(\Delta^+ + 1)$ -approximable, where $\Delta^+ = \max\{|N_{in}^+(u)| : u \in V\}$.*

We remark that the result in Theorem 2 only depends on the topology of G^+ , no matter which is the topology of G^- . Throughout the next subsection we shall follow the opposite way, that is, we shall study the complexity of 1-USTSS in relation to the topology of G^- , no matter which is the topology of G^+ .

5.1. 1-USTSS and G^-

Let $G = (V, A^+ \cup A^-)$ be a signed graph. Notice that any $I \subseteq V$ is a target set for G only if, for each $u \in I$, it holds that $N_{out}^-(u) \subseteq I$. In turn, if I is a target set for G containing u and, hence, $N_{out}^-(u)$, then I must contain also $N_{out}^-(v)$, for any $v \in N_{out}^-(u)$. More precisely, if we define, for each $u \in V$, the set $\mathcal{E}(u)$ containing u and all the nodes reachable from u by a path in G^- , that is,

$$\mathcal{E}(u) = \{u\} \cup \{v \in V : u \in \pi^-(v)\},$$

then the next fact holds, which follows from the previous observation by a simple inductive argument.

Fact 1: *If $I \subseteq V$ is a target set for $G = (V, A^+ \cup A^-)$ then, for any $u \in I$, $\mathcal{E}(u) \subseteq I$.*

Observe now that, if G^- is strongly connected, then, for any $u \in V$, it holds that $\mathcal{E}(u) = V$; hence, by Fact 1, we get

Fact 2: *If G^- is strongly connected, then the only target set for G (no matter about its latency) is V .*

For any $E \subseteq V$, we define the set

$$S(E) = E \cup \{u \in V : d^+(E, u) = 1\} = E \cup \{u \in V : \exists z \in E \cap N_{in}^+(u)\},$$

that is, the set of nodes that are or may become informed within time step 1 whenever all nodes in E are contained in the set of the initially informed nodes.

Lemma 5. *$I \subseteq V$ is a 1-target set for $G = (V, A^+ \cup A^-)$ if and only if $S(I) = V$ and, for any $u \in I$, $\mathcal{E}(u) \subseteq I$.*

Proof. If $I \subseteq V$ is a 1-target set for $G = (V, A^+ \cup A^-)$, then, trivially, $S(I) = V$ and, by Fact 1, for any $u \in I, \mathcal{E}(u) \subseteq I$.

Conversely, suppose that $S(I) = V$ and, for any $u \in I, \mathcal{E}(u) \subseteq I$. Then, for any node $x \in V - I$,

- since $S(I) = V$, there exists $y \in I$ such that $(y, x) \in A^+$ (that is, y may inform x at time step 1) and, furthermore,
- since for any $u \in I, \mathcal{E}(u) \subseteq I$, then, for any $z \in V$ such that $(z, x) \in A^-$, since $x \in \pi^-(z)$ and $x \notin I, z \notin I$ (that is, any node that could forbid x to become informed is not in I). \square

Lemma 5 allows us to prove membership to PO of 1-USTSS when restricted to chain and generalized chain graphs. The chain graph over the node set $V = \{v_1, \dots, v_n\}$ is made of the set of arcs $\{(v_j, v_{j-1}) : j = 2, \dots, n\}$. A generalized chain graph consists of a set C_1, \dots, C_n of strongly connected components linked to each other in a chain-like topology, that is, for any $j = 2, \dots, n - 1$ and for any $u \in C_j$, arc (u, v) exists only if $v \in C_{j-1}$.

Theorem 3. *The 1-USTSS problem restricted to instances $G = (V, A^+ \cup A^-)$ such that $G^- = (V, A^-)$ is a chain or a generalized chain graph is in PO.*

Proof. If G^- is a chain, that is, $V = \{v_1, \dots, v_n\}$ and $A^- = \{(v_j, v_{j-1}) : j = 2, \dots, n\}$, then $\mathcal{E}(v_1) = \{v_1\}, \mathcal{E}(v_2) = \{v_1, v_2\} = \{v_2\} \cup \mathcal{E}(v_1)$ and, in general, for $i \geq 2, \mathcal{E}(v_i) = \{v_i\} \cup \mathcal{E}(v_{i-1})$. Hence, $\mathcal{E}(v_1) \subseteq \dots \subseteq \mathcal{E}(v_n)$.

Correspondingly, $S(\mathcal{E}(v_1)) \subseteq \dots \subseteq S(\mathcal{E}(v_n))$: indeed, for any $i = 1, \dots, n - 1$,

$$S(\mathcal{E}(v_i)) = \mathcal{E}(v_i) \cup \{u \in V : \exists z \in \mathcal{E}(v_i) \cap N_{in}^+(u)\} \subseteq \mathcal{E}(v_{i+1}) \cup \{u \in V : \exists z \in \mathcal{E}(v_{i+1}) \cap N_{in}^+(u)\}.$$

As a consequence, by Lemma 5, $\mathcal{E}(v_m)$ is a minimum size 1-target set for G , where $m = \min\{j \in \{1, \dots, n\} : S(\mathcal{E}(v_j)) = V\}$.

Let G^- be a generalized chain and let C_1, \dots, C_n be its strongly connected components. As already stated in the observation leading to Fact 2, for any $j = 1, \dots, n$ and for any $u, v \in C_j$, it holds that $\mathcal{E}(u) = \mathcal{E}(v)$: hence, in what follows we shall use the notation \mathcal{E}_j to refer to any set $\mathcal{E}(u)$ for $u \in C_j$. By the same argument holding for chain graphs, for any $j = 1, \dots, n$, we get that \mathcal{E}_j contains all nodes in any C_i with $i \leq j$. Hence, again by a similar reasoning to that holding in the chain case, since $\mathcal{E}_1 \subseteq \dots \subseteq \mathcal{E}_n$, we get that \mathcal{E}_m is a minimum size 1-target set for G , where $m = \min\{j \in \{1, \dots, n\} : S(\mathcal{E}_j) = V\}$. \square

Lemma 5 shows that any 1-target set is actually obtained by choosing a subcollection of sets in the collection $\mathcal{E} = \{\mathcal{E}(u) : u \in V\}$. We now introduce the collection $\bar{\mathcal{E}}$ as the closure of \mathcal{E} under non-empty intersections: that is, $\mathcal{E} \subseteq \bar{\mathcal{E}}$, and, for each $E_1, E_2 \in \bar{\mathcal{E}}$ such that $E_1 \cap E_2 \neq \emptyset, E_1 \cup E_2 \in \bar{\mathcal{E}}$.

The introduction of $\bar{\mathcal{E}}$ allows us to describe any target set for G as the union of pairwise disjoint subsets, that is:

Fact 3: *For any target set I for $G = (V, A^+ \cup A^-)$, there exists $\mathcal{B} \subseteq \bar{\mathcal{E}}$ such that $I = \cup_{B \in \mathcal{B}} B$ and, for any $B_1, B_2 \in \mathcal{B}, B_1 \cap B_2 = \emptyset$.*

Based on fact 3, we now show how to get in over polynomial time a $(1 + \log |V|)$ -approximate solution for 1-USTSS by a (non-polynomial-time) reduction from 1-USTSS to MINIMUM WEIGHTED SET COVER (WSC) and by the $(1 + \log |V|)$ -approximation algorithm for WSC [45]. Later in this section, we shall discuss cases in which the procedure overall runs in polynomial time.

Recall that an instance of WSC is a triple $\langle X, \mathcal{S}, w \rangle$ where $\mathcal{S} \subseteq 2^X, \cup_{S \in \mathcal{S}} S = X$, and $w : \mathcal{S} \rightarrow \mathbb{N}^+$; the goal is finding a minimum weight set $\mathcal{C} \subseteq \mathcal{S}$ such that $\cup_{C \in \mathcal{C}} C = X$, where the weight of \mathcal{C} is defined as $w(\mathcal{C}) = \sum_{C \in \mathcal{C}} w(C)$.

Let $G = (V, A^+ \cup A^-)$ be a signed graph; the corresponding instance $\langle X_G, \mathcal{S}_G, w_G \rangle$ of WSC is obtained by setting $X_G = V, \mathcal{S}_G = \{S(E) : E \in \bar{\mathcal{E}}\}$ and, for each $E \in \bar{\mathcal{E}}, w_G(S(E)) = |E|$.

Lemma 6. Let $G = (V, A^+ \cup A^-)$ be a signed graph and $E_1, \dots, E_k \in \bar{\mathcal{E}}$; $I_0 = \cup_{j=1}^k E_j$, is a 1-target set for G if and only if $\mathcal{C} = \{S(E_1), \dots, S(E_k)\}$ is a set cover for X_G . Furthermore, if E_1, \dots, E_k are pairwise disjoint then $|I_0| = w_G(\mathcal{C})$.

Proof. Let E_1, \dots, E_k be elements of $\bar{\mathcal{E}}$ and let $I_0 = \cup_{j=1}^k E_j$. By definition, I_0 is a 1-target set for G if and only if $S(I_0) = V$. Hence,

$$S(I_0) = \bigcup_{j=1}^k S(E_j) = V = X_G,$$

that is, $\mathcal{C} = \{S(E_1), \dots, S(E_k)\}$ is a set cover of X .

If E_1, \dots, E_k are pairwise disjoint, then

$$|I_0| = \left| \bigcup_{j=1}^k E_j \right| = \sum_{j=1}^k |E_j| = \sum_{j=1}^k w_G(S(E_j)) = w_G(\mathcal{C}).$$

□

The previous lemma shows that the proposed transformation is actually a reduction and that, if a 1-target set for G is described as the union of pairwise disjoint elements of $\bar{\mathcal{E}}$, then its size equals the weight of the corresponding set cover of X_G .

Notice that, if \mathcal{C}^* is an optimal solution for the instance $\langle X_G, \mathcal{S}_G, w_G \rangle$ of WSC corresponding to G then, for any $S(E_1), S(E_2) \in \mathcal{C}^*$, it holds that $E_1 \cap E_2 = \emptyset$: indeed, if there exist $S(E_1), S(E_2) \in \mathcal{C}$ such that $E_1 \cap E_2 \neq \emptyset$, since $E_1 \cup E_2 \in \bar{\mathcal{E}}$ and, thus, $S(E_1 \cup E_2) \in \mathcal{S}_G$ and, moreover, $S(E_1) \cup S(E_2) = S(E_1 \cup E_2)$, then the set $\mathcal{C}' = \mathcal{C}^* - \{S(E_1), S(E_2)\} \cup \{S(E_1 \cup E_2)\}$ would still be a set cover for X_G and

$$\begin{aligned} \sum_{\mathcal{C} \in \mathcal{C}'} w_G(\mathcal{C}) &= \sum_{\mathcal{C} \in \mathcal{C}^*} w_G(\mathcal{C}) - w_G(S(E_1)) - w_G(S(E_2)) + w_G(S(E_1 \cup E_2)) \\ &= \sum_{\mathcal{C} \in \mathcal{C}^*} w_G(\mathcal{C}) - |E_1| - |E_2| + |E_1 \cup E_2| < \sum_{\mathcal{C} \in \mathcal{C}^*} w_G(\mathcal{C}), \end{aligned}$$

so contradicting that \mathcal{C}^* is an optimal solution for $\langle X_G, \mathcal{S}_G, w_G \rangle$.

As a consequence of the previous reasoning and of Lemma 6, if \mathcal{C}^* is an optimal solution for the instance $\langle X_G, \mathcal{S}_G, w_G \rangle$ of WSC corresponding to G then the set $I^* = \{E \in \bar{\mathcal{E}} : S(E) \in \mathcal{C}^*\}$ is an optimal 1-target set for G .

Consider now the following procedure to compute a 1-target set of a signed graph $G = (V, A^+ \cup A^-)$: first compute set $\bar{\mathcal{E}}$, then the instance $\langle X, \mathcal{S}, w \rangle$ of WSC and, finally apply to it Chvatal's algorithm in [45] to get a solution $\mathcal{C} = \{S(E_1), \dots, S(E_k)\}$ such that $w_G(\mathcal{C}) \leq (1 + \log |X_G|)w_G(\mathcal{C}^*)$, where \mathcal{C}^* is an optimal solution for $\langle X_G, \mathcal{S}_G, w_G \rangle$.

Needless to say, it may well happen that the sets E_1, \dots, E_k are not pairwise disjoint, so that the second part of Lemma 6 cannot be exploited to bound the size of the so found 1-target set. In this case, we derive from \mathcal{C} a new set cover $\mathcal{C}' = \{S(E'_1), \dots, S(E'_h)\}$ for X_G such that $w_G(\mathcal{C}') \leq w_G(\mathcal{C})$ and E'_1, \dots, E'_h are pairwise disjoint: set $\mathcal{C}' = \mathcal{C}$ and then, while there exist $S(E), S(F) \in \mathcal{C}'$ such that $E \cap F \neq \emptyset$, replace in \mathcal{C}' the pair $S(E), S(F)$ by $S(E \cup F)$. Notice that, at any replacement, \mathcal{C}' is still a set cover for $\langle X_G, \mathcal{S}_G, w_G \rangle$ and its weight does not increase since $w_G(S(E \cup F)) - (w_G(S(E)) + w_G(S(F))) = |E \cup F| - (|E| + |F|) \leq 0$.

Hence, in $poly(|G|, |\bar{\mathcal{E}}|)$ -time we have computed a 1-target set $I_0 = \{E'_1, \dots, E'_h\}$ for G such that

$$|I_0| = w_G(\mathcal{C}') \leq w_G(\mathcal{C}) \leq (1 + \log |X_G|)w_G(\mathcal{C}^*) = (1 + \log |V|)|I^*|.$$

This proves the following theorem.

Theorem 4. *It is possible to find an $(1 + \log |V|)$ -approximate solution for an instance $G = (V, A^+ \cup A^-)$ of 1-USTSS in $\text{poly}(|G|, |\bar{\mathcal{E}}|)$ -time.*

In general, the algorithm described so far does not run in polynomial time since the size of $\bar{\mathcal{E}}$ is not polynomially bounded by $|G|$. Actually, the size of $\bar{\mathcal{E}}$ is strongly related to the topology of G^- and in what follows we shall consider topologies yielding families of polynomial size. Without loss of generality, we shall only consider (weakly) connected graphs.

A source tree is a weakly connected directed graph containing one source node s connected to each other node of the graph by exactly one path and such that, for any pair of nodes u and v , if there is a path from u to v then such a path is contained in the path from s to v . Similarly as for the generalized chains, a generalized source tree consists of a set of strongly connected components linked to each other in a source tree-like topology. This means that, if we replace each strongly connected component in G^- by a single node, we get a source tree.

Corollary 1. *1-USTSS is $(1 + \log |V|)$ -approximable for instances $G = (V, A^+ \cup A^-)$ such that G^- is a source tree or a generalized source tree.*

Proof. When the graph G^- is a (directed) source tree then, for each $u \in V$, $\mathcal{E}(u)$ is the set of nodes in the subtree of G^- rooted at u . Hence, if $\mathcal{E}(u) \cap \mathcal{E}(v) \neq \emptyset$ then either $\mathcal{E}(u) \subseteq \mathcal{E}(v)$ or $\mathcal{E}(v) \subseteq \mathcal{E}(u)$. This proves that $\bar{\mathcal{E}} = \{\mathcal{E}(u) \mid u \in V\}$, that is, $|\bar{\mathcal{E}}| = |V|$.

A similar reasoning to that used in Theorem 3 for the case of the generalized chain allows to extend to generalized trees the result about trees. \square

6. Conclusions and Open Problems

This paper aims at studying the impact of negative relations (witnessing antagonism or distrusting) to diffusion processes within a network. We first defined a diffusion rule, generalizing the Linear Threshold Model to signed relations, and then we studied the complexity of computing a minimum size target set with respect to it. We finally considered a set of restrictions under which the problem is approximable to some extent.

Our approximability results all refer to 1-target sets. We proved that the lower bound $(\log |V| + 1)$ on the approximation ratio is tight for a simple topology of the graph of the negative relations, that is, the generalized source tree, and that our problem is in PO if such graph is a generalized chain. Needless to say, broadening the set of topologies of the graph of the negative relations still yielding reasonable approximation is an interesting research direction. After the results in [3,24], studying the problem restricted to graphs of bounded tree-width or clique-width looks like a worthwhile research direction.

After the discussion in Section 2, we remark that 1-target sets correspond to some kind of dominating sets, that we have called dominating sets with constrained pairs: given a graph $G = (V, E)$ and a set $C \subseteq V \times V$, any node $v \in V$ is dominated by a set $D \subseteq V$ if $v \in D$ or v has a neighbor in D and, for each $(u, v) \in C$, $u \notin D$. We have, thus, proved that the bound $(\log |V| + 1)$ on the approximation ratio of the DOMINATING SET problem still holds for our constrained problem when (V, C) is a generalized source tree. Again, broadening the set of topologies of the graph of the negative relations still yielding $O(\log |V|)$ -approximation seems an interesting research direction. Finally, our $(\Delta^+ + 1)$ -approximation ratio is exponentially far from the $[\log(\Delta + 1) + 1]$ -approximation ratio holding for the MINIMUM DOMINATING SET problem [43]. Narrowing the gap is an issue deserving attention too.

Author Contributions: Conceptualization, M.D.I. and G.V.; Methodology, M.D.I. and G.V.; Formal analysis, M.D.I. and G.V.; Writing—original draft preparation, M.D.I. and G.V.; Writing—review and editing, M.D.I. and G.V.. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Domingos, P.; Richardson, M. Mining the Network Value of Customers. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 26 August 2001; pp. 57–66.
2. Richardson, M.; Domingos, P. Mining Knowledge-sharing Sites for Viral Marketing. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, AB, Canada, 23 July 2002; pp. 61–70.
3. Ben-Zwi, O.; Hermelin, D.; Lokshtanov, D.; Newman, I. Treewidth governs the complexity of target set selection. *Discret. Optim.* **2011**, *8*, 87–96. [[CrossRef](#)]
4. Chen, W.; Wang, C.; Wang, Y. Scalable Influence Maximization for Prevalent Viral Marketing in Large-scale Social Networks. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 25 July 2010; pp. 1029–1038.
5. Granovetter, M. Threshold Models of Collective Behavior. *Am. J. Sociol.* **1978**, *6*, 1420–1443. [[CrossRef](#)]
6. Kempe, D.; Kleinberg, J.; Tardos, É. Maximizing the Spread of Influence Through a Social Network. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 24 August 2003; pp. 137–146.
7. Kempe, D.; Kleinberg, J.; Tardos, É. Influential Nodes in a Diffusion Model for Social Networks. In *Automata, Languages and Programming*; Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M., Eds.; Springer: Berlin, Germany, 2005; Volume 3580, pp. 1127–1138.
8. Leskovec, J.; Adamic, L.A.; Huberman, B.A. The Dynamics of Viral Marketing. *ACM Trans. Web* **2007**, *1*, 5-es. [[CrossRef](#)]
9. Contagion, M.S. *The Review of Economic Studies*; Oxford University Press: New York, NY, USA, 2000; pp. 57–78.
10. Ackerman, E.; Ben-Zwi, O.; Wolfowitz, G. Combinatorial Model and Bounds for Target Set Selection. *Theor. Comput. Sci.* **2010**, *411*, 4017–4022. [[CrossRef](#)]
11. Bazgan, C.; Chopin, M.; Nichterlein, A.; Sikora, F. Parameterized approximability of maximizing the spread of influence in networks. *J. Discret. Algorithms* **2014**, *27*, 54–65. [[CrossRef](#)]
12. Centeno, C.C.; Dourado, M.C.; Penso, L.D.; Rautenbach, D.; Szwarcfiter, J.L. Irreversible conversion of graphs. *Theor. Comput. Sci.* **2011**, *412*, 3693–3700. [[CrossRef](#)]
13. Chen, N. On the approximability of influence in social networks. *SIAM J. Discret. Math.* **2009**, *23*, 1400–1415. [[CrossRef](#)]
14. Chiang, C.Y.; Huang, L.H.; Li, B.J.; Wu, J.; Yeh, H.G. Some results on the target set selection problem. *J. Comb. Optim.* **2013**, *25*, 702–715. [[CrossRef](#)]
15. Chiang, C.Y.; Huang, L.H.; Yeh, H.G. Target Set Selection Problem for Honeycomb Networks. *SIAM J. Discret. Math.* **2013**, *27*, 310–328. [[CrossRef](#)]
16. Chopin, M.; Nichterlein, A.; Niedermeier, R.; Weller, M. Constant Thresholds Can Make Target Set Selection Tractable. *Theor. Comput. Syst.* **2014**, *55*, 61–83. [[CrossRef](#)]
17. Zaker, M. On dynamic monopolies of graphs with general thresholds. *Discret. Math.* **2012**, *312*, 1136–1143. [[CrossRef](#)]
18. Bettencourt, L.M.A.; Cintron-Arias, A.; Kaiser, D.I.; Castillo-Chavez, C. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Phys. Stat. Mech. Appl.* **2006**, *364*, 513–536. [[CrossRef](#)]
19. Nekovee, M.; Moreno, Y.; Bianconi, G.; Marsili, M. Theory of rumour spreading in complex social networks. *Phys. Stat. Mech. Appl.* **2007**, *374*, 467–470. [[CrossRef](#)]

20. Goldenberg, J.; Libai, B.; Muller, E. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Market. Lett.* **2001**, *12*(3), 211–223. [[CrossRef](#)]
21. Flocchini, P.; Lodi, E.; Luccio, F.; Pagli, L.; Santoro, N. Irreversible dynamos in tori. In *Euro-Par'98 Parallel Processing*; Pritchard, D., Reeve, J., Eds.; Springer: Berlin, Germany, 1998; Volume 1470, pp. 554–562.
22. Luccio, F.; Pagli, L.; Sanossian, H. Irreversible dynamos in butterflies. In *Proceedings of the 6th Colloquium on Structural Information and Communication Complexity*, Lacanau-Ocean, France, 1–3 July 1999; pp. 204–218.
23. Peleg, D. Size bounds for dynamic monopolies. *Discret. Appl. Math.* **1998**, *86*, 263–273. [[CrossRef](#)]
24. Cicalese, F.; Cordasco, G.; Gargano, L.; Milanič, M.; Vaccaro, U. Latency-Bounded Target Set Selection in Social Networks. *Theor. Comput. Sci.* **2014**, *535*, 1–15. [[CrossRef](#)]
25. Dreyer Jr, P.A.; Roberts, F.S. Irreversible k-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discret. Appl. Math.* **2009**, *157*, 1615–1627. [[CrossRef](#)]
26. Gargano, L.; Hell, P.; Peters, J.G.; Vaccaro, U. Influence Diffusion in Social Networks under Time Window Constraints. *Theor. Comput. Sci.* **2015**, *584*, 53–66. [[CrossRef](#)]
27. Dreyer, P. A. Applications and Variations of Domination in Graphs. Ph.D. Thesis, Rutgers University, Camden, NJ, USA, 2000.
28. Alba, J.W.; Hutchinson, J.W.; Lynch J.G. Memory and decision making. In *Handbook of Consumer Behavior*; Robertson, T.S., Kassarian H., Eds.; Prentice-Hall, Englewood Cliffs, NJ, USA, 1991; pp. 1–49.
29. Chen, Y.; Iyer, G.; Pazgal, A. Limited Memory, Categorization and Competition. *Market. Sci.* **2010**, *29*, 650–670. [[CrossRef](#)]
30. Chen, W.; Collins, A.; Cummings, R.; Ke, T.; Liu, Z.; Rincon, D.; Sun, X.; Wang, Y.; Wei, W.; Yuan, Y. Influence maximization in social networks when negative opinions may emerge and propagate. In *Proceedings of the 2011 SIAM International Conference on Data Mining*, Mesa, AZ, USA, 28–30 April 2011; pp. 379–390.
31. Nazemian, A.; Taghiyareh, F. Influence maximization in independent cascade model with positive and negative word of mouth. In *Proceedings of the 6th International Symposium on Telecommunications (IST)*, Tehran, Iran, 6–8 November 2012; pp. 854–860.
32. Stich, L.; Golla, G.; Nanopoulos, A. Modelling the spread of negative word-of-mouth in online social networks. *J. Decis. Syst.* **2014**, *23*, 203–221. [[CrossRef](#)]
33. Yang, S.; Wang, S.; Truong, V.A. Online Learning and Optimization Under a New Linear-Threshold Model with Negative Influence. *arXiv* **2019**, arXiv:1911.03276. Available online: <https://arxiv.org/pdf/1911.03276.pdf> (accessed on 27 January 2020).
34. Ahmed, S.; Ezeife, C.I. Discovering influential nodes from trust network. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, New York, NY, USA, 18 March 2013; pp. 121–128.
35. Li, Y.; Chen, W.; Wang, Y.; Zhang, Z.L. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, Rome, Italy, 4 February 2013; pp. 657–666.
36. Easley, D.; Kleinberg, J. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. *Significance* **2012**, *9*, 43–44.
37. Galhotra, S.; Arora, A.; Roy, S. Holistic influence maximization: Combining scalability and efficiency with opinion-aware models. In *Proceedings of the 2016 International Conference on Management of Data*, San Francisco, CA, USA, 14 June 2016; pp. 743–758.
38. Hosseini-Pozveh, M.; Zamanifar, K.; Naghsh-Nilchi, A.R.; Dolog, P. Maximizing the spread of positive influence in signed social networks. *Intell. Data Anal.* **2016**, *20*, 199–218. [[CrossRef](#)]
39. Baumeister, R.F.; Bratslavsky, E.; Finkenauer, C. Bad is stronger than good. *Rev. Gen. Psychol.* **2001**, *5*, 323–370. [[CrossRef](#)]
40. Peeters, G.; Czapinski, J. Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *Eur. Rev. Soc. Psychol.* **1990**, *1*, 33–60. [[CrossRef](#)]
41. Rozin, P.; Royzman, E.B. Negativity bias, negativity dominance, and contagion. *Personal. Soc. Psychol. Rev.* **2001**, *5*, 296–320. [[CrossRef](#)]
42. Taylor, S.E. Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychol. Bull.* **1991**, *110*, 67–85. [[CrossRef](#)]

43. Foerster, K.T. Approximating Fault-Tolerant Domination in General Graphs. In Proceedings of the Tenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), New Orleans, LA, USA, 6 January 2013.
44. Khanna, S.; Sudan, M.; Trevisan, L.; Williamson, D.P. The approximability of constraint satisfaction problems. *SIAM J. Comput.* **2000**, *30*(6), 1863–1920. [[CrossRef](#)]
45. Chvátal, V. A greedy heuristic for the set covering problem. *Math. Oper. Res.* **1979**, *4*, 233–235. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).