On the Inapproximability of Broadcasting Time (Extended Abstract)

Christian Schindelhauer*

International Computer Science Institute, Berkeley, USA
Med. Universität zu Lübeck, Institut für Theoretische Informatik, Lübeck, Germany
schindel@tcs.mu-luebeck.de

Abstract. We investigate the problem of broadcasting information in a given undirected network. At the beginning information is given at some processors, called sources. Within each time unit step every informed processor can inform only one neighboring processor. The broadcasting problem is to determine the length of the shortest broadcasting schedule for a network, called the *broadcasting time* of the network.
We show that there is no efficient approximation algorithm for the broadcasting time of a network with a single source unless $\mathcal{P} = \mathcal{NP}$. More formally, it is $\mathcal{NP}$-hard to distinguish between graphs $G = (V, E)$ with broadcasting time smaller than $b \in \Theta(\sqrt{|V|})$ and larger than $(\frac{57}{56} - \epsilon)b$ for any $\epsilon > 0$.
For ternary graphs it is $\mathcal{NP}$-hard to decide whether the broadcasting time is $b \in \Theta(\log |V|)$ or $b + \Theta(\sqrt{b})$ in the case of multiples sources. For ternary networks with single sources, it is $\mathcal{NP}$-hard to distinguish between graphs with broadcasting time smaller than $b \in \Theta(\sqrt{|V|})$ and larger than $b + c\sqrt{\log b}$.
We prove these statements by polynomial time reductions from E3-SAT.
Classification: Computational complexity, inapproximability, network communication.

1 Introduction

Broadcasting reflects the sequential and parallel aspects of disseminating information in a network. At the beginning the information is available only at some *sources*. The goal is to inform all nodes of the given network. Every node may inform another neighboring node after a certain *switching time*. Along the edges there may be a *delay*, too. Throughout this abstract the switching time is one time unit and edges do not delay information. This model is called *Telephone model* and represents the broadcasting model in its original setting [GaJo79].

The restriction of the broadcasting problem to only one information source v_0 has often been considered, here called *single source broadcasting problem* **(SB)**. Note that the broadcasting time $\boldsymbol{b(G, v_0)}$ is at least $\log_2 |V|$ for a graph $G = (V, E)$, since during each round the number of informed vertices can at most

* Parts of this work are supported by a stipend of the "Gemeinsames Hochschulsonderprogramm III von Bund und Länder" through the DAAD.

K. Jansen and S. Khuller (Eds.): APPROX 2000, LNCS 1913, pp. 226–237, 2000.

double. The smallest graph providing this lower bound is a *binomial tree* F_n [HHL88]: F_0 consists of a single node and F_{n+1} consists of disjunct subtrees $F_0, \ldots, F_n$ whose roots $r_0, \ldots, r_n$ are connected to the new root r_{n+1}. Also the *hyper-cube* $C_n = \{\{0,1\}^n\}, \{\{w0v, w1v\} \mid w, v \in \{0,1\}^*\}$ has this minimum broadcasting time since binomial trees can be derived by deleting edges.

The upper bound on $b(G)$ is $|V| - 1$, which is needed for the chain graph representing maximum sequential delay (Fig. 1) and the star graph (Fig. 2) producing maximum parallel delay. The topology of the processor network highly influences the broadcasting time and much effort was given to the question how to design networks optimized for broadcasting, see [LP88,BHLP92,HHL88].

Throughout this paper the communication network and the information sources are given and the task is to find an efficient broadcasting schedule. The original problem deals with *single sources* and its decision problem, called **SBD**, to decide whether the broadcasting time is less or equal a given deadline T_0, is $\mathcal{NP}$-complete [GaJo79,SCH81]. Slater et al. also show, for the special case of trees, that a divide-and-conquer strategy leads to a linear time algorithm. This result can be generalized for graphs with a small tree-width according to a tree decomposition of the edges [JRS98]. However, SBD remains $\mathcal{NP}$-complete even for the restricted case of ternary planar graphs or ternary graphs with logarithmic depth [JRS98].

Bar-Noy et al. [BGNS98] present a polynomial-time approximation algorithm for the *single source broadcasting problem* **(SB)** with an approximation factor of $O(\log |V|)$ for a graph $G = (V, E)$. SB is approximable within $O(\frac{\log |V|}{\log \log |V|})$ if the graph has bounded tree-width with respect to the standard tree decomposition [MRSR95].

Adding more information sources leads to the *multiple source broadcasting problem* **(MB)**. It is known to be NP-complete even for constant broadcasting time, like 3 [JRS98] or 2 [Midd93]. This paper solves the open problem whether there are graphs that have a non-constant gap between the broadcasting time $b(G)$ and a polynomial time computable upper bound. In [BGNS98] this question was solved for the more general *multicast* model proving an inapproximability factor bound of $3 - \epsilon$ for any $\epsilon > 0$. In this model switching time and edge delay may differ for each node and instead of the whole network a specified sub-network has to be informed.

It was an open problem whether this lower bound could be transfered to the *Telephone* model. In this paper, we solve this problem using a polynomial time reduction from E3-SAT to SB. The essential idea makes use of the high degree of the reduction graph's source. A good broadcasting strategy has to make most of its choices there and we show that this is equivalent to assigning variables of an E3-CNF-formula. A careful book-keeping of the broadcasting times of certain nodes representing literals and clauses gives the lower bound of $\frac{57}{56} - \epsilon$.

We show for ternary graphs and multiple sources that graphs with a broadcasting time $b \in \Theta(\log |V|)$ cannot be distinguished from those with broadcasting time $b + c\sqrt{b}$ for some constant c. This result implies that it is $\mathcal{NP}$-hard to distinguish between ternary graphs with the single source broadcasting time of $b \in \Theta(\sqrt{|V|})$ and graphs with broadcasting time $b + c\sqrt{\log b}$.

The paper is organized as follows. In Section 2 formal notations are introduced, in the next section the general lower bound of SB is proved. We present in section 4 lower bounds for the ternary case. Section 5 concludes and summarizes these results.

2 Notation

Edges of the given undirected graph may be directed to indicate the information flow along an edge.

Definition 1. *Let $G = (V, E)$ be an undirected graph with a set of vertices $V_0 \subseteq V$, called the* **sources**. *The task is to compute the* **broadcasting time** $\boldsymbol{b(G, V_0)}$*, the minimum length T of a* **broadcast schedule** *S. This is a sequence of sets of directed edges $S = (E_1, E_2, \dots, E_{T-1}, E_T)$. Their nodes are in the sets $V_0, V_1, \dots, V_T = V$, where for $i > 0$ we define $V_i := V_{i-1} \cup \{v \mid (u, v) \in E_i \text{ and } u \in V_{i-1}\}$. A broadcast schedule S fulfills the properties*

1. *$E_i \subseteq \{ (u, v) \mid u \in V_{i-1}, \{u, v\} \in E \}$ and*
2. *$\forall u \in V_{i-1} : \mid E_i \cap (\{u\} \times V)| \leq 1$.*

The set of nodes V_i has received the broadcast information by round i. For an optimal schedule with length T, the set V_T is the first to include all nodes of the network. E_i is the set of edges used for sending information at round i. Each processor $u \in V_{i-1}$ can use at most one of its outgoing edges in every round.

Definition 2. *Let S be a broadcast schedule for (G, V_0), where $G = (V, E)$. The* **broadcasting time of a node** *$v \in V$ is defined as* $\boldsymbol{b_S(v)} = \min\{i \mid v \in V_i\}$. *A broadcast schedule S is called* **busy** *if the following holds.*

1. *$\forall \{v, w\} \in E \ : \ b_S(w) > b_S(v) + 1 \quad \Rightarrow \quad \exists w' \in V \ : \ (v, w') \in E_{b_S(w)-1}$*
2. *$\forall v \in V \setminus \{v_0\} \ : \ |\bigcup_i E_i \ \cap (V \times \{u\})| = 1$.*

In a busy broadcasting schedule, every processor tries to inform a neighbor in every step starting from the moment it is informed. When this fails it stops. By this time, all its neighbors are informed. Furthermore, every node is informed only once. Every schedule can be transformed into a busy schedule within polynomial time without increasing the broadcasting time of any node. From now on, every schedule is considered to be busy. In [BGNS98] this argument is generalized (the authors call busy schedules *not lazy*).

A chain is defined by $C_n = (\{v_1, \dots, v_n\}, \ \{\{v_i, v_{i+1}\}\})$ (Fig. 1), and a star by $S_n = (\{v_1, \dots, v_n\}, \{\{v_1, v_i\} \mid i > 1\})$ (Fig. 2).

Fact 1. *There is only one busy broadcast strategy that informs a chain with k interior nodes. Let its ends v, w be informed in time $b_w - k \leq b_v \leq b_w$. Then the chain is informed in time $\lceil (b_v + b_w + k)/2 \rceil$ assuming that the ends have no obligations for informing other nodes.*

There are $n!$ busy broadcast schedules for the star S_n that describe all permutations of $\{1, \dots, n\}$ by $(b_S(v_1), \dots, b_S(v_n))$.

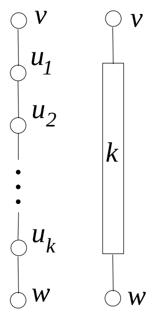

Fig. 1. The chain and its symbol.

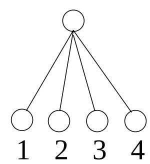

Fig. 2. The star and a busy broadcasting schedule.

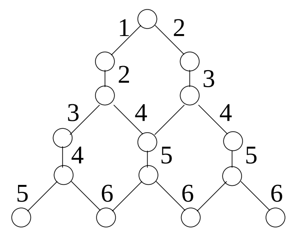

Fig. 3. The ternary pyramid and a busy schedule.

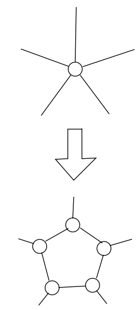

Fig. 4. A ring of n nodes as replacement of a star-sub-graph.

3 The General Lower Bound

This section presents a polynomial time reduction from E3-SAT to SB and the proof of the constant inapproximability factor. E3-SAT denotes the satisfiability problem of Boolean CNF-formulas with exactly three literals in each clause.

Theorem 1 *[Håst97]. For any $\epsilon > 0$ it is $\mathcal{NP}$-hard to distinguish satisfiable E3-SAT formulas from E3-SAT formulas for which only a fraction $7/8+\epsilon$ of the clauses can be satisfied, unless $\mathcal{P} = \mathcal{NP}$.*

Let F be a 3-CNF with m clauses $c_1, \ldots, c_m$ and variables $x_1, \ldots, x_n$. Let $a(i)$ denote the number of occurrences of the positive literal x_i in F. It is possible to assume that every variable occurs as often positive as negated in F, since in the proof of Theorem 1 this property is fulfilled. Let $\delta := 2\ell m'$, where $m' := \sum_{i=1}^{n} a(i)$ with ℓ being a large number to be chosen later on. Note that $m = \frac{2}{3}m'$.

The formula F is reduced to an undirected graph $G_{F,\ell}$ (see Fig. 5). The source v_0 and its δ neighbors $x^b_{i,j,k}$ form a star S_δ ($b \in \{0,1\}$, $i \in \{1,\ldots,n\}$, $j \in \{1,\ldots,a(i)\}$, $k \in \{1,\ldots,\ell\}$). We call the nodes $x^b_{i,j,k}$ *literal nodes.* They belong to ℓ disjunct isomorphic subgraphs $G_1, \ldots, G_\ell$. A subgraph G_k contains literal nodes $x^b_{i,j,k}$, representing the literal x^b_i ($x^1_i = x_i$, $x^0_i = \overline{x_i}$).

As a basic tool for the construction of a sub-graph G_k, a chain $C_p(v,w)$ is used starting at nodes v and ending at w with p interior nodes that are not incident to any other edge of the graph. Between the literal nodes corresponding with a variable x_i in G_k we insert chains $C_\delta(x^0_{i,j,k}, x^1_{i,j',k})$ for all $i \in \{1,\ldots,n\}$ and $j, j' \in \{1,\ldots,a(i)\}$.

For every clause $c_\nu = x^{b_1}_{i_1} \vee x^{b_2}_{i_2} \vee x^{b_3}_{i_3}$ we insert *clause nodes* $c_{\nu,k}$ which we connected via three chains $C_{\delta/2}(c_{\nu,k}, x^{b_\rho}_{i_\rho,j_\rho,k})$ for $\rho \in \{1,2,3\}$ of length $\delta/2$ to their corresponding literal nodes $x^{b_1}_{i_1,j_1,k}, x^{b_2}_{i_2,j_2,k}, x^{b_3}_{i_3,j_3,k}$. This way every literal node is connected to one clause node. This completes the construction of G_k.

The main idea of the construction is that the assignment of a variable x_i indicates when the corresponding literal nodes have to be informed.

Lemma 1. *If F is satisfiable, then $b(G_{F,\ell}, v_0) \leq \delta + 2m' + 2$.*

Proof: The busy schedule S informs all literal nodes directly by v_0. Let $\alpha_1, \ldots, \alpha_n$ be a satisfying assignment of F. The literal nodes $x^{\alpha_i}_{i,j,k}$ of graph G_k are informed within the time period $(k-1)m'+1, \ldots, km'$. The literal nodes $x^{\overline{\alpha_i}}_{i,j,k}$ are informed within the time period $\delta - km' + 1, \ldots, \delta - (k-1)m'$.

Note that m' is a trivial upper bound for the degree at a literal node. So, the chains between two literal nodes can be informed in time $\delta + 2m' + 1$. A clause node can be informed in time $km' + \delta/2 + 1$ by an assigned literal node of the first type, which always exists since $\alpha_1, \ldots, \alpha_n$ satisfies F. Note that all literal nodes corresponding to the second type are informed within $\delta - (k-1)m'$. So the chains between those and the clause node are informed in time $\delta + 2m' + 2$. ∎

Lemma 2. *Let S be a busy broadcasting schedule for $G_{F,\ell}$. Then,*

1. *every literal node will be informed directly from the source v_0, and*
2. *for $c_{\nu,k} = x^{\alpha_1}_{i_1,j_1,k} \vee x^{\alpha_2}_{i_2,j_2,k} \vee x^{\alpha_3}_{i_3,j_3,k}$: $b_S(c_{\nu,k}) > \frac{\delta}{2} + \min_\rho\{b_S(x^{\alpha_\rho}_{i_\rho,j_\rho,k})\}$.*

Proof:

1. Every path between two literal nodes that avoids v_0 has at least length $\delta+1$. By Fact 1 even the first informed literal node has no way to inform any other literal node before time point δ, which is the last time a literal node is going to be informed by v_0.
2. follows by 1.

∎

If only one clause per Boolean formula is not satisfied, this lemma implies that if F is not satisfiable, then $b(G_{F,\ell}, \{v_0\}) > \delta + \ell$. A better bound can be achieved if the inapproximability result of Theorem 1 is applied. A busy schedule S for graph $G_{F,\ell}$ defines an assignment for F. Then, we categorize every literal as *high*, *low* or *neutral*, depending on the consistency of the time of information. Clause nodes are classified either as *high* or *neutral*. Every unsatisfied clause of the E3-SAT-formula F will increase the number of high literals. Besides this, high and low literal nodes come in pairs, yet possibly in different subgraphs G_k and $G_{k'}$. The overall number of the high nodes will be larger than those of the low nodes.

Theorem 2. *For every $\epsilon > 0$ there exist graphs $G = (V, E)$ with broadcasting time at most $b \in \Theta(\sqrt{|V|})$ such that it is $\mathcal{NP}$-hard to distinguish those from graphs with broadcasting time at least $(\frac{57}{56} - \epsilon)b$.*

Proof: Consider an unsatisfiable E3-SAT-formula F, the above described graph $G_{F,\ell}$ and a busy broadcasting schedule S on it. The schedule defines for each subgraph G_k an assignment $x_{1,k}, \ldots, x_{n,k} \in \{0,1\}^n$ as follows. Assign the variable $x_{i,k} = \alpha$ if the number of delayed literal nodes with $b_S(x^\alpha_{i,j,k}) > \delta/2$ is smaller than those with $b_S(x^{\overline{\alpha}}_{i,j,k}) > \delta/2$. If both numbers are equal, w.l.o.g. let $x_{i,k} = 0$.

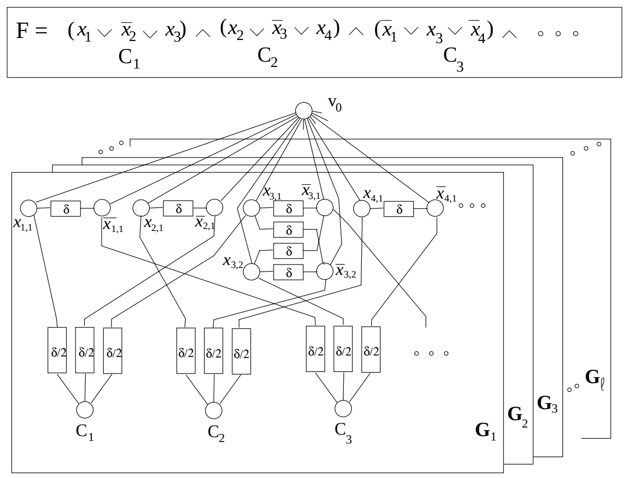

Fig. 5. The reduction graph $G_{F,\ell}$.

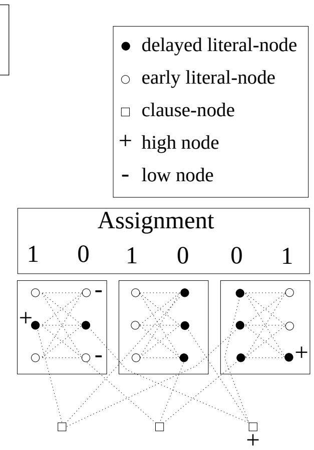

Fig. 6. High and low literal nodes.

1. A literal node $c^{\alpha}_{i,j,k}$ is **coherently assigned**, iff $b_S(c^{\alpha}_{i,j,k}) \leq \delta/2 \;\Leftrightarrow\; x_{i,k} = \alpha$. All coherently assigned literal nodes are **neutral**.
2. A literal node $x^{\alpha}_{i,j,k}$ is **high** if it is not coherently assigned and delayed, i.e. $x_{i,k} = \alpha$ and $b_S(x^{\alpha}_{i,j,k}) > \delta/2$.
3. A literal node $x^{\alpha}_{i,j}$ is **low** if it is not coherently assigned and not delayed, i.e. $x_{i,k} = \overline{\alpha}$ and $b_S(x^{\alpha}_{i,j,k}) \leq \delta/2$.
4. A clause node $c_{\nu,k}$ is **high**, if all its three connected literal nodes are coherent and delayed, i.e. $\forall \rho \in \{1,2,3\}\;\; b_S(x^{\alpha_i}_{i_\rho,j_\rho,k}) > \delta/2$.
5. All other clause nodes are **neutral**.

Every high literal node with broadcasting time $\delta/2+\epsilon_1$ for $\epsilon_1 > 0$ can be matched to a neutral delayed literal node $x^{\overline{\alpha}}_{i,j',k}$ with broadcasting time $b_S(x^{\overline{\alpha}}_{i,j',k}) = \delta/2 + \epsilon_2$ for $\epsilon_2 > 0$. Fact 1 shows that the chain between both of them can be informed in time $\delta + \frac{\epsilon_1+\epsilon_2}{2}$ at the earliest.

For a high clause node with literal nodes $x^{\alpha_i}_{i_\rho,j_\rho,k}$ and broadcasting times $b_S(x^{\alpha_i}_{i_\rho,j_\rho,k}) = \delta/2 + \epsilon_\rho$ with $\epsilon_1, \epsilon_2, \epsilon_3 > 0$, Lemma 2 shows that this high clause node gets the information not earlier than $\delta + \min\{\epsilon_1, \epsilon_2, \epsilon_3\}$. So, the chain to the most delayed literal node will be informed at $\delta + (\min\{\epsilon_1, \epsilon_2, \epsilon_3\} + \max\{\epsilon_1, \epsilon_2, \epsilon_3\})/2$ at the earliest.

Lemma 3. *Let q be the number of low literal nodes, p the number of high literal nodes, and p' the number of high clause nodes. Then the following holds:*

1. $p = q$,
2. $b_S(G_{F,\ell}, v_0) \;\geq\; \delta + p$,
3. $b_S(G_{F,\ell}, v_0) \;\geq\; \delta + (p + 3p')/2$.

Proof:

1. Consider the set of nodes $x^{\alpha}_{i,j,k}$, for $j \in \{1,\ldots,a(i)\}$ and $\alpha \in \{0,1\}$. For this set let $p_{i,k}$ be the number of high nodes, $q_{i,k}$ the number of low nodes and $r_{i,k}$ the number of nodes with time greater than $\delta/2$. By the definition of high and low nodes the following holds for all $i \in \{1,\ldots,n\}$, $k \in \{1,\ldots,\ell\}$:

$$r_{i,k} - p_{i,k} + q_{i,k} = a(i) \ .$$

Fact 1 and Lemma 2 show that half of the literal nodes are informed within $\delta/2$ and the rest later on:

$$\sum_{i,k} r_{i,k} = \delta/2 = \sum_{i,k} a(i) \ ,$$

It then it follows that:

$$q - p = \sum_{i,k} r_{i,k} - p_{i,k} + q_{i,k} - a(i) = 0 \ .$$

2. Note that we can match each of the p high (delayed) literal node $x^{\alpha}_{i,j,k}$ to a coherent delayed literal node $x^{\overline{\alpha}}_{i,j',k}$. Furthermore, these nodes have to inform a chain of length δ. If the latest of the high nodes and its partners is informed at time $\delta/2+\epsilon$, then Fact 1 shows that the chain cannot be informed earlier than $\delta + \epsilon/2$.
The broadcasting time of all literal nodes is different. Therefore it holds $\epsilon \geq 2p$, proving $b_S(G_{F,\ell}, v_0) \ \geq \ \delta + p$.
3. Every high clause node is connected to three neutral delayed literal nodes. The task to inform all chains to the three literal nodes is done at time $\delta+\epsilon'/2$ at the earliest, if $\delta/2 + \epsilon'$ is the broadcasting time of the latest literal node. For p' high clause nodes, there are $3p'$ corresponding delayed neutral literal nodes. Furthermore, there are p delayed high literal nodes (whose matched partners may intersect with the $3p'$ neutral literal nodes). Nevertheless, the latest high literal node with broadcasting time $\delta/2 + \epsilon''$ causes a broadcast time on the chain to a neutral delayed literal node of at least $\delta + \epsilon''/2$.
From both groups consider the most delayed literal node $v_{\max}$. Since every literal node has a different broadcasting time it holds that $\epsilon'' \geq 3p' + p$, and thus $b_S(v_{\max}) \geq \delta + (3p' + p)/2$. ∎

Suppose all clauses are satisfiable. Then Lemma 1 gives an upper bound for the optimal broadcasting time of $b(G_{F,\ell}, v_0) \ \leq \ \delta + 2m' + 2$.

Let us assume that at least κm of the m clauses are unsatisfied for every assignment. Consider a clause node that represents an unsatisfied clause with respect to the assignment which is induced by the broadcast schedule. Then at least one of the following cases can be observed:

- The clause node is high, i.e. its three literal nodes are coherently assigned.
- The clause node is neutral and one of its three literal nodes is low.
- The clause node is neutral and one of its three literal nodes is high.

Since each literal node is chained to one clause node only, this implies

$$\kappa\ell m \ \leq \ p' + p + q \ = \ p' + 2p \ .$$

The case $p \geq 3p'$ implies $p \geq \frac{3}{7}(2p+p')$. Then it holds for the broadcasting time of any busy schedule S:

$$b_S(G_{F,\ell}, v_0) \ \geq \ \delta + p \ \geq \ \delta + \tfrac{3}{7}(p' + 2p) \ .$$

Otherwise, if $p < 3p'$, then $\frac{1}{2}(p + 3p') \geq \frac{3}{7}(2p + p')$ and

$$b_S(G_{F,\ell}, v_0) \ \geq \ \delta + \tfrac{1}{2}(p + 3p') \ \geq \ \delta + \tfrac{3}{7}(p' + 2p) \ .$$

Note that $\delta = 3m\ell$. Combining both cases, it follows that

$$b_S(G_{F,\ell}, v_0) \ \geq \ \delta + \tfrac{3}{7}\kappa\ell m \ = \ \delta\left(1 + \tfrac{1}{7}\kappa\right) \ .$$

For any $\epsilon > 0$ this gives, choosing $\ell \in \Theta(m)$ for sufficient large m

$$\frac{b_S(G_{F,\ell}, v_0)}{b(G_{F,\ell}, v_0)} \ \geq \ \frac{1 + \frac{1}{7}\kappa}{1 + \frac{2m'+2}{\delta}} \ \geq \ 1 + \tfrac{1}{7}\kappa - \epsilon \ .$$

Theorem 1 states $\kappa = \frac{1}{8} - \epsilon''$ for any $\epsilon'' > 0$ which implies claimed lower bound of $\frac{57}{56} - \tilde{\epsilon}$ for any $\tilde{\epsilon} > 0$. Note that the number of nodes of $G_{F,\ell}$ is in $\Theta(m^4)$ and $\delta \in \Theta(m^2)$. ∎

4 Inapproximability Results for Ternary Graphs

The previous reduction used graphs $G_{F,\ell}$ with a large degree at the source node. To address ternary graphs with multiple sources we modify this reduction as follows.

The proof uses a reduction from the E3-SAT-6 problem: a CNF formula with n variables and $m = n/2$ clauses is given. Every clause contains exactly three literals and every variable appears three times positive and three times negative, but does not appear in a clause more than once. The output is the maximum number of clauses that can be satisfied simultaneously by some assignment to the variables.

Lemma 4. *For some $\epsilon > 0$, it is $\mathcal{NP}$-hard to distinguish between satisfiable 3CNF-6 formulas, and 3CNF-6 formulas in which at most a $(1-\epsilon)$-fraction of the clauses can be satisfied simultaneously.*

Proof: Similar as Proposition 2.1.2 in [Feig98]. Here, every second occurrence of a variable is replaced with a fresh variable when reducing from E3-SAT. This way the number of positive and negative literals remains equally high. ∎

How can the star at the source be replaced by a ternary sub-graph that produces high differences between the broadcasting times of the literal nodes? It turns out that a good way to generate such differences in a very symmetric setting is a complete binary tree. Using trees instead of a star complicates the situation. A busy broadcasting schedule informs $\binom{d}{t}$ leaves in time $d+t$ where in the star graph only one was informed in time t. This is the reason for the dramatic decrease of the inapproximability bound.

The ternary reduction graph $G'_{F,\ell}$, given a 3CNF-6-formula F and a number ℓ to be chosen later, consists of the following sub-graphs (see Fig. 7).

1. The sources $v_1, \dots, v_n$ are roots of *complete binary trees* $B_1, \dots, B_n$ with depth $\delta = \log(12\ell)$ and leaves $v^i_1, \dots, v^i_{2^\delta}$. ℓ will be chosen such that δ is an even number.

A constant fraction of the leaves of B_i are the literal nodes $x^\alpha_{i,j,k}$ of a subgraph G_k. The rest of them, $y^\alpha_{i,j}$ is connected in pairs via δ-chains. For an accurate description we introduce the following definitions.

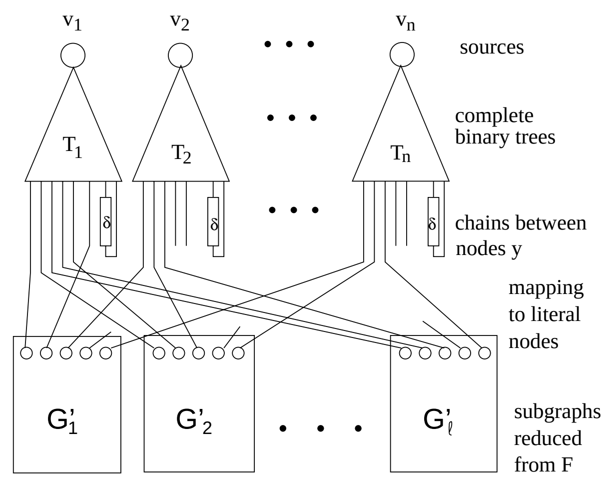

Fig. 7. The reduction graph $G'_{F,\ell}$.

Let $f_\delta(p,\delta) := \sum_{i=\delta/2+1}^{\delta/2+p} \binom{\delta}{i}$. Since $\binom{\delta}{\delta/2} \le \frac{2^\delta}{\sqrt{\delta}}$ and $\binom{\delta}{\delta/2+\sqrt{\delta}} \ge \frac{2^\delta}{10\sqrt{\delta}}$ it holds for $p \in \{1, \dots, \sqrt{\delta}\}$: $\frac{p}{10}\frac{2^\delta}{\sqrt{\delta}} \le f_\delta(p) \le p\frac{2^\delta}{\sqrt{\delta}}$. For $g_d(x) := \min\{p \mid f(p,d) \ge x\}$ this implies for $x \in [0, \frac{2^\delta}{10}]$: $x\frac{\sqrt{\delta}}{2^\delta} \le g_\delta(x) \le 10x\frac{\sqrt{\delta}}{2^\delta}$. Note that f_δ and g_δ are monotone increasing.

Every node of B_i is labeled by a binary string. If r is the root, **label(*r*)** is the empty string λ. The two successing nodes v_1, v_2 of a node w are labeled by **label(*w*)0** and **label(*w*)1** Two leaves x, y are called *opposite* if label(x) can be derived from label(y) by negating every bit. For a binary string let $\Delta(s) := |\#_1(s) - \#_0(s)|$ be the difference of occurrences of 1 and 0 in s. Consider an *indexing* $v^i_1, \dots, v^i_{2^\delta}$ of the leaves of B_i such that for all $j \in \{1, \dots, 2^\delta - 1\}$: $\Delta(\text{label}(v^i_j)) \le \Delta(\text{label}(v^i_{j+1}))$, and v^i_j and $v^i_{2^\delta-j+1}$ have opposite labels for all $j \in \{1, \dots, 2^\delta\}$.

2. For every binary tree B_i according to these indices the literal nodes of G_k are defined by $x^0_{i,j,k} = v^i_{2^{\delta-1}+3(k-1)+j}$ and $x^1_{i,j,k} = v^i_{2^{\delta-1}-3(k-1)-j+1}$ for $j \in \{1,\ldots,3\}$, and $k \in \{1,\ldots,\ell\}$.
3. The other leaves of B_i are connected pairwise by chains of length δ such that opposite leaves of a tree represent free literal nodes $y^0_{i,j}$ and $y^1_{i,j}$. These nodes are not part of any sub-graph G_k.
4. The sub-graphs G_k for $k \in \{1,\ldots,k\}$ described in the previous section have a degree 5 at the literal nodes. These nodes are replaced with rings of size 5 to achieve degree 3 (see Fig. 4).

Theorem 3. *It is $\mathcal{NP}-$hard to distinguish ternary graphs $G=(V,E)$ with multiple sources and broadcasting time $b \in \Theta(\log|V|)$ from those with broadcasting time $b+c\sqrt{b}$ for any constant c.*

Proof Sketch: If F is satisfiable, then there is a coherently assigning broadcast schedule with $b(G'_{F,\ell}) \leq 2\delta+4$.

An analogous observation to Lemma 2 for a busy broadcasting schedule S for $G'_{F,\ell}$ is the following.

1. Every literal node will be informed directly from the source of its tree;
2. For all $i \in \{1,\ldots,n\}$ and for all $t \in \{0,\ldots,\delta\}$ it holds
$$|\{j \in \{1,\ldots,2^\delta\} \mid b_S(v^i_j) = t+\delta\}| \;=\; \binom{\delta}{t}\;;$$
3. For $c_{\nu,k} = x^{\alpha_1}_{i_1,j_1,k} \vee x^{\alpha_2}_{i_2,j_2,k} \vee x^{\alpha_3}_{i_3,j_3,k}$: $b_S(c_{\nu,k}) = \frac{\delta}{2} + \min_\rho\{b_S(x^{\alpha_\rho}_{i_\rho,j_\rho,k})\} + O(1)$.

Again literal nodes are defined to be either low, high, or neutral. Clause nodes are either high or neutral. For the number q of low literals, p of high literals, and p' the number of high clauses it holds $p=q$. There are $2p$, resp. $3p'$ nodes in different chains that are informed later than $2\delta-1$. Therefore there is a tree B_k that is involved in the delayed information of $2p/n$, resp. $3p'/n$ nodes. Using g_δ it is possible to describe a lower bound of the time delay caused by B_k as follows.

$$b_S(G_{F,\ell}) \;\geq\; 2\delta - 1 + \frac{1}{2}\max\left\{g_\delta\left(\frac{2p}{n}\right), g_\delta\left(\frac{3p'}{n}\right)\right\} \;.$$

Let us assume that at least κm clauses are unsatisfied for every assignment. The constant fraction of y-leaves of trees T_i can be seen as an additional set of unused literal nodes. Now consider a clause node that represents an unsatisfied clause with respect to the assignment which is induced by the broadcast schedule. Then there is at least a high clause node, a neutral clause node connected to a low literal node, or a neutral clause node connected to a high literal node.

Since each literal node is chained to at most one clause node, this implies

$$\kappa\ell m \;\leq\; p' + p + q \;=\; p' + 2p\;.$$

Note that $24\ell \geq 2^\delta$. The observations above now imply

$$b_S(G'_{F,\ell}, \{v_1,\ldots,v_n\}) \;\geq\; 2\delta - 1 + \frac{1}{2} g_\delta\!\left(\frac{\kappa\ell m}{2n}\right) \;\geq\; 2\delta + \epsilon\sqrt{\delta}$$

for some $\epsilon > 0$. Since for the set of nodes V of $G'_{F,\ell}$ it holds $|V| \in \Theta(\ell m \log \ell)$ it is sufficient to choose ℓ as a non constant polynomial of m. ∎

Theorem 4. *It is $\mathcal{NP}$ − hard to distinguish ternary graphs $G = (V, E)$ with* **single sources** *and broadcasting time $b \in \Theta(\sqrt{|V|})$ from those with broadcasting time $b + c\sqrt{\log b}$ for some constant c.*

Proof: We start to combine the reduction graph of the preceding theorem with a ternary pyramid (see Fig 3). The single source v_0 is the top of the pyramid. The n leaves have been previously the sources. Note that the additional amount of broadcasting time in a pyramid is $2n$ for $n-1$ nodes and $2n-1$ for one node for any busy broadcasting schedule. Thus, the former sources are informed nearly at the same time.

For the choice $\ell \in \Theta(\frac{m}{\log m})$ the number of nodes of the new graph is bounded by $\Theta(m^2)$. The broadcasting time increases from $\Theta(\log m)$ of $G'_{F,\ell}$ to $\Theta(m)$ and the indistinguishable difference remains $\Theta(\sqrt{\log m})$. ∎

5 Conclusions

The complexity of broadcasting time is a key for understanding the obstacles to efficient communication in networks. This article answers the open question stated most recently in [BGNS98], whether single source broadcasting in the Telephone model can be approximated within any constant factor. Until now, the best upper bound approximation ratio for broadcasting time is known $O(\log |V|)$ [BGNS98] and the lower bound was known as one additive time unit. Thus, a lower constant bound of a factor of $\frac{57}{56} - \epsilon$ is a step forward. Yet there is room for improvement.

It is possible to transfer this result to bounded degree graphs. But the reconstruction of sub-graphs with large degree decrease the lower bound dramatically. Nevertheless, this paper improves on the inapproximability ratio in the single source case up to $1 + \Theta\left(\sqrt{\frac{\log |V|}{|V|}}\right)$, instead of $1 + 1/\Theta(\sqrt{|V|})$ known so far [JRS98]. The upper bound for approximating the broadcasting time of a ternary graph is a constant factor. So matching upper and lower bounds remain unknown.

From a practical point of view, network structures are often uncertain because of dynamic and unpredictable changes. And if the network is static, it is hardly ever possible to determine the ratio between switching time on a single processor and the delay on communication links. But if these parameters are known for every processor and communication link it turns out that an inapproximability factor $3 - \epsilon$ applies [BGNS98]. For the simplest timing model, the Telephone model, this paper shows that developing a good broadcasting strategy is also a computationally infeasible task.

Acknowledgment

I would like to thank Andreas Jakoby and Rüdiger Reischuk for encouraging me to revisit broadcasting problems on which had worked together before. Furthermore, I thank some anonymous referees for their hints to the literature and Antonio Piccolboni for his remarks.

References

BGNS98. A. Bar-Noy, S. Guha, J. Naor, B. Schieber, *Multicasting in heterogeneous networks*, In Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing, 1998, 448-453.

BHLP92. J.-C. Bermond, P. Hell, A. Liestman, and J. Peters, *Broadcasting in Bounded Degree Graphs*, SIAM J. Disc. Math. 5, 1992, 10-24.

Feig98. U. Feige, *A threshold of* $\ln n$ *for approximating set cover*, Journal of the ACM, Vol. 45, 1998 (4):634-652.

GaJo79. M. Garey and D. Johnson, *Computers and Intractability, A Guide To the Theory of* $\mathcal{NP}$*-Completeness*, Freeman 1979.

Håst97. J. Håstad, *Some optimal inapproximability results,* In Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing, 1997, 1-10.

HHL88. S. Hedetniemi, S. Hedetniemi, and A. Liestman, *A Survey of Gossiping and Broadcasting in Communication Networks*, Networks 18, 1988, 319-349.

JRS98. A. Jakoby, R. Reischuk, C. Schindelhauer, *The Complexity of Broadcasting in Planar and Decomposable Graphs,* Discrete Applied Mathematics 83, 1998, 179-206.

LP88. A. Liestman and J. Peters, *Broadcast Networks of Bounded Degree*, SIAM J. Disc. Math. 4, 1988, 531-540.

Midd93. M. Middendorf, *Minimum Broadcast Time is* $\mathcal{NP}$*–complete for 3–regular planar graphs and deadline 2*, Information Processing Letters, 46, 1993, 281-287.

MRSR95. M. V. Marathe, R. Ravi, R. Sundaram, S.S. Ravi, D.J. Rosenkrantz, H.B. Hunt III, *Bicriteria network design problems*, Proc. 22nd Int. Colloquium on Automata, Languages and Programming, Lecture Notes in Comput. Sci. 944, Springer-Verlag, 1995, 487-498.

SCH81. P. Slater, E. Cockayne, and S. Hedetniemi, *Information Dissemination in Trees*, SIAM J. Comput. 10, 1981, 692-701.