Journal of Complex Networks (2016) **4**, 177–186
doi:10.1093/comnet/cnv020
Advance Access publication on 21 July 2015

Network installation under convex costs

Alexander Gutfraind[†]
Division of Epidemiology and Biostatistics, University of Illinois at Chicago, Chicago, IL, USA
[†]Corresponding author. Email: agutfrai@uic.edu

and

Jeremy Kun, Ádám D. Lelkes and Lev Reyzin
Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL, USA

Edited by: Cristopher Moore

[Received on 12 November 2014; accepted on 15 June 2015]

We study the Neighbour-Aided Network Installation Problem (NANIP) introduced previously which asks for a minimal cost ordering of the nodes of a graph, where the cost of visiting a node is a function of the number of its neighbours that have already been visited. This problem has applications in resource management and disaster recovery. In this paper, we analyse the computational hardness of NANIP. In particular we show that this problem is NP-hard even when restricted to convex decreasing cost functions, give a linear approximation lower bound for the greedy algorithm, and prove a general sub-constant approximation lower bound. Then we give a new integer programming formulation of NANIP and empirically observe its speedup over the original integer programme.

Keywords: infrastructure network; disaster recovery; permutation optimization; neighbour-aided network installation problem.

1. Introduction

We motivate our study with an example from infrastructure networks. It is well known that many vital infrastructure systems can be represented as networks, including transport, communication and power networks. Large parts of these networks can be severely damaged in the event of a natural disaster. When faced with large-scale damage, authorities must develop a plan for restoring the networks. A particularly challenging aspect of the recovery is the lack of infrastructure, such as roads or power, necessary to support the recovery operations. For example, to clear and rebuild roads, equipment must be brought in, but many of the access roads are themselves blocked and damaged. Abstractly, as the recovery progresses, previously recovered nodes provide resources that help reduce the cost of rebuilding their neighbours. We call this phenomenon 'neighbour aid'.

Recently, [1] introduced and analysed a simple model of neighbour aided recovery in terms of a convex discrete optimization problem called the *Neighbour-Aided Network Installation Problem* (NANIP). We will henceforth use the terms 'recover', 'visit' and 'install' interchangeably. For simplicity, we assume that during the recovery of a network all of its nodes and edges must be visited and restored. They asked how to optimize the recovery schedule in order to minimize the total cost? This is also the question we address herein.

In the NANIP problem, the cost of recovering a node depends only on the number of its already recovered neighbours, capturing the intuition that neighbour aid is the determining factor of the cost of rebuilding a new node. NANIP offers a stylized model for disaster recovery of networks (among other applications) but the interest in disaster recovery of networks is not new. A partial list of existing studies include [2–7]. A common framework is to consider infrastructure systems as a set of interdependent network flows, and formulate the problem of minimizing the cost of repairing such damaged networks. Another class of models [8] develops a stochastic optimization problem for stockpiling resources and then distributing them following a disaster. More abstract problems related to NANIP are the single processor scheduling problem [9], the linear ordering problem [10] and the study of tournaments in graph theory [11].

NANIP assumes that certain tasks are dependent and cannot be performed in parallel, but unlike many scheduling problems, there are no partial order constraints. Like the travelling salesman problem (TSP) [12], the NANIP problem also asks for an optimal permutation of the vertices of the graph but, unlike in the case of the TSP, the cost associated with visiting a given node could depend on all of the nodes visited before the given node. Another key difference between NANIP and TSP is that in NANIP it is allowed to visit nodes that are not neighbours of any previously visited nodes. As we will see, such disconnected traversals provide $\Omega(\log(n))$ multiplicative improvements over connected ones.

Since neighbour aid is assumed to reduce the cost of recovery, we are mainly interested in decreasing cost functions. Furthermore, since convexity for decreasing functions captures the 'law of diminishing returns', i.e. that as the number of recovered neighbours increases, the per-node value of the aid provided by one neighbour decreases, convex decreasing functions are of special interest. Although [1] gave NP-hardness of NANIP for general cost via a straightforward reduction from Maximum Independent Set, the cost function used there was increasing, thus leaving the complexity of the convex decreasing case an open question. In this paper, we show this problem is NP-hard as well. We also provide a new convex integer programming (IP) formulation and analyse the performance of the greedy algorithm, showing that its worst case approximation ratio is $\Theta(n)$.

2. Preliminaries

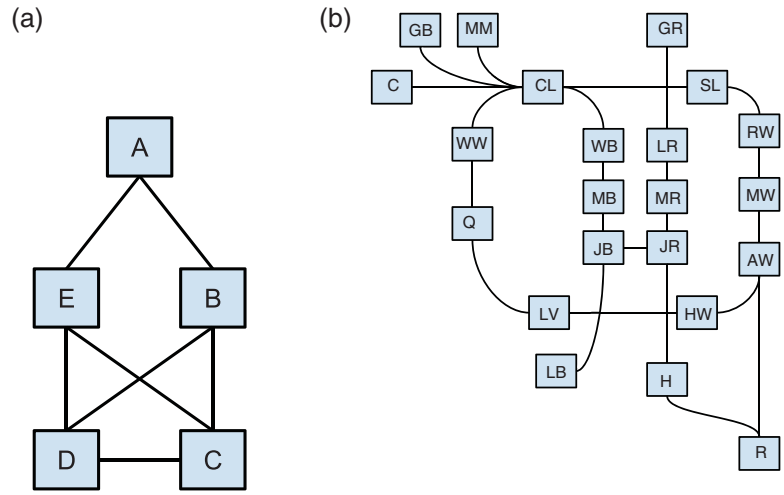
An instance of NANIP is specified by an undirected graph $G = (V, E)$ and a real-valued function $f : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$. The function f represents the cost of installing a vertex v, where the argument is the number of neighbours of v that have already been installed. Hence, the domain of f is the non-negative integers, bounded by the maximum degree of G (for terminology see [11]). The goal is to find a permutation of the nodes that minimizes the total cost of the network installation. The cost of installing node $v_t \in V$ under a permutation σ of V is given by

$$f(r(v_t, G, \sigma)),$$

where $r(v_t, G, \sigma)$ is the number of nodes adjacent to v_t in G that appear before v_t in the permutation σ. The total cost of installing G according to σ is given by

$$C_G(\sigma) = \sum_{t=1}^{n} f(r(v_t, G, \sigma)). \tag{1}$$

The problem is illustrated in Fig. 1. Generally, the choice of f depends on the application, and f will often be convex decreasing.



Simple NANIP instance Central Chicago "Loop"

FIG. 1. Illustrations of NANIP. (a) Simple instance. When $f(0) = 2$, $f(1) = 1$ and $f(k) = 0$ for $k \geqslant 2$, the naive sequence $\sigma = (A, B, C, D, E)$ gives cost of $4 = 2 + 1 + 1 + 0 + 0$, but all optimal solutions (such as (D, C, B, E, A)) have cost 3. (b) Actual metro stations and their connections in downtown Chicago 'Loop'. With the same f, any optimal sequences must visit CL station before at least one of its neighbours.

We assume that G is connected and undirected, unless we note otherwise. If G has multiple connected components, NANIP could be solved on each component independently without affecting the total cost.

We begin by quoting a preliminary lemma from [1] which establishes that all the arguments used in calculating the node costs must sum to m, the number of edges in the network.

LEMMA 1 (1) For any network G, and any permutation σ of the nodes of G,

$$\sum_{t=1}^{n} r(v_t, G, \sigma) = m. \tag{2}$$

One application of this lemma is the case of a linear cost function $f(k) = ak + b$, for some real numbers a and b. With such a function the optimization problem is trivial in that all installation permutations have the same cost.

In the next section, we will prove hardness results about NANIP; let us recall some relevant definitions.

DEFINITION 2 An optimization problem is called *strongly NP-hard* if it is NP-hard and the optimal value is a positive integer bounded by a polynomial of the input size.

DEFINITION 3 An algorithm is an *efficient polynomial time approximation scheme (EPTAS)* for an optimization problem if, given a problem instance and an approximation factor ε, it runs in time $O(F(\varepsilon)n^c)$ for some constant c and some function F and finds a solution whose objective value is within an ε fraction of the optimum. An EPTAS is called a *fully polynomial time approximation scheme (FPTAS)* it runs in polynomial in the size of the problem instance and $1/\varepsilon$.

A strongly NP-hard optimization problem cannot have an FPTAS unless P=NP: otherwise, if n denotes the input size and p denotes the polynomial such that the optimum value is bounded by $p(n)$, setting $\varepsilon = 1/2p(n)$ for the FPTAS would yield an exact polynomial time algorithm.

Some NP-hard problems become efficiently solvable if a natural parameter is fixed to some constant. Such problems are called fixed parameter tractable.

DEFINITION 4 FPT, the set of *fixed parameter tractable* problems, is the set of languages L of the form $\langle x, k\rangle$ such that there is an algorithm running in time $O(F(k)n^c)$ for some function F and constant c deciding whether $\langle x, k\rangle \in L$.

An example of a fixed parameter tractable problem is the vertex cover problem (where the parameter is the size of the vertex cover). Problems believed to be fixed parameter intractable include the graph colouring problem (the parameter being the number of colours) and the clique problem (with the size of the clique as parameter).

For parametrized languages, there is a natural fixed parameter tractable analogue of polynomial time reductions. These so-called *fpt-reductions* are used to define hardness for classes of parametrized languages, similarly to how NP-hardness is defined using polynomial time reductions. One important class of parametrized languages is $W[1]$. For the definition of $W[1]$ and for more background on parametrized complexity, we refer the reader to the monograph of Downey and Fellows [13]. They proved that under standard complexity-theoretic assumptions, $W[1]$ is a strict superset of FPT; consequently, $W[1]$-hard problems are fixed parameter intractable. We will use this fact to show the fixed parameter intractability of NANIP.

3. Convex decreasing NANIP is NP-hard

We now consider the hardness of solving NANIP with convex decreasing cost functions.

THEOREM 5 The NANIP is strongly NP-hard when f is convex decreasing; as a consequence it admits no FPTAS.

Proof. We reduce from CLIQUE, that is, the problem of deciding given a graph $G = (V, E)$ whether it contains as an induced subgraph the complete graph on k vertices. Given a graph $G = (V, E)$ with $n = |V|$ and an integer k, we construct an instance of NANIP on a graph G' with a convex cost function $f(i)$ as follows. Define G' by adding k new vertices $u_1, \ldots, u_k$ to G which are made adjacent to every vertex in V but not to each other, establishing an independent set of size k. Define the cost function

$$f(i) = f_k(i) = \begin{cases} k - i & \text{if } i \leqslant k \\ 0 & \text{otherwise.} \end{cases}$$

Let $M = \sum_{i=0}^{k} f(i) = k(k+1)/2$. In a traversal σ whose first k vertices yield cost M, every new vertex must be adjacent to every previously visited vertex, i.e. the vertices form a k-clique. Moreover, M is the lower bound on the cost incurred by the first k vertices of any traversal of G'.

Suppose that G has a clique of size k, and denote by $v_1, \ldots, v_k$ the vertices of the clique, with $v_{k+1}, \ldots, v_n$ the remaining vertices of G. Then the following ordering is a traversal of G' of cost exactly M:

$$v_1, \ldots, v_k, u_1, \ldots, u_k, v_{k+1}, \ldots, v_n.$$

Conversely, let $w_1, \ldots, w_{n+k}$ be an ordering of the vertices of G' achieving cost M. Then by the above, the vertices $w_1, \ldots, w_k$ must form a k-clique in G'. In the case these k prefix vertices are all vertices of G we are done. Otherwise, the independence of the u_i's implies that at most one u_i is used in $w_1, \ldots, w_{k+1}$; using more would incur a total cost greater than M. In this case the $k-1$ remaining vertices of the prefix form a $(k-1)$-clique of G. Since it is NP-hard to approximate CLIQUE within a polynomial factor [14], this proves the NP-hardness of convex decreasing NANIP.

Moreover, since the optimum value of a NANIP instance obtained by this reduction is at most k^2 which is upper bounded by n^2, the size of the NANIP instance, it also follows that convex decreasing NANIP is strongly NP-hard and therefore does not admit an FPTAS. □

The cost function $f_k(i)$ used in the proof of Theorem 5 is parametrized by k. Call NANIP$_k$ the subproblem of NANIP with cost functions of finite support where the size of the support is k. Because we consider NANIP$_k$ a subproblem of general NANIP, stronger parametrized hardness results for the former give insights about the latter. Indeed, the following corollary is immediate.

COROLLARY 6 NANIP$_k$ is $W[1]$-hard.

Proof. CLIQUE is $W[1]$-complete when parametrized by the size of the clique. $W[1]$-hardness is preserved by so-called *fpt*-reductions (see [13]), and the reduction from the proof of Theorem 5 is such a reduction. □

In particular, standard complexity assumptions imply from this that NANIP$_k$ is not fixed-parameter tractable and has no EPTAS. Now we will show that the same reduction can be used to obtain a stronger approximation lower bound of $(1 + n^{-c})$ for all $c > 0$. First a lemma.

LEMMA 7 Let G' and f constructed as in the proof of Theorem 5, and let σ denote a (not necessarily optimal) NANIP traversal. Suppose V denote the vertices of G and U denote the vertices of the independent set. If σ' is obtained from σ by moving the U to positions $k+1, \ldots, 2k$ (without changing the precedence relations of the vertices in V), then $C_{G'}(\sigma') \leqslant C_{G'}(\sigma)$.

Proof. Consider the positions in σ of the first k vertices from G, and let $i_1, \ldots, i_k$ be the positions of the vertices from U. Call $u_1 = \sigma(i_1), \ldots, u_k = \sigma(i_k)$.

Case 1: $i_1 > k$. In this case, all the u_i are free (since they are all connected to $\sigma(1), \ldots, \sigma(k)$), as are all vertices visited after $\sigma(i_k)$). If $i_1 > k+1$, apply the cyclic permutation $\gamma_1 = (k+1, k+2, \ldots, i_1)$ to move u_1 to position $k+1$. The cost of visiting u_1 is still zero, and the cost of the other manipulated vertices does not increase because they each gain one previously visited neighbour. Now repeat this manipulation with $\gamma_s = (k+s, k+s+1, \ldots, i_s)$ for $s = 2, \ldots, k$. An identical argument shows the cost never increases, and at the end we have precisely σ'.

Case 2: $i_1 \leqslant k$. In this case u_1 is not free since it has less than k installed neighbours. By moving some nodes from V before the nodes U, we will make the u_i's free without increasing the cost of the moved nodes too much, and thus reduce the problem to Case 1. Let j be the index of the first $v \in V$ that occurs after i_1. Apply the cyclic permutation $\xi = (i_1, i_1+1, \ldots, j)$ to move v before u_1. The cost of v increases by at most $j - i_1$ (and this is not tight since it is possible that $j > k+1$). But since all $\sigma(i_1), \sigma(i_1+1), \ldots, \sigma(j-1) \in U$, and they each gain a neighbour as a result of applying ξ, so their total cost decreases by exactly $j - i_1$, and the total cost of σ does not increase. Now repeatedly apply ξ (using the new values of i_1, j) until $i_1 = k+1$. Then apply case 1 to finish. □

THEOREM 8 For all $c > 0$, there is no efficient $(1 + n^{-c})$-approximation algorithm for NANIP on graphs with n vertices with convex decreasing cost functions, unless P = NP.

Proof. It is NP-hard to distinguish a clique number of at least 2^R from a clique number of at most $2^{\delta R}$ in graphs on $2^{(1+\delta)R}$ vertices ($\delta > 0$) [14]. We will reduce this problem to finding an $(1 + n^{-c})$-approximation for NANIP. In particular, we will show that there is no efficient C-approximation algorithm for NANIP, where

$$C = \frac{k}{k+1}\left(1 + \frac{1}{k^{2\varepsilon}}\right)$$

and $k = n^{1/(1+\delta)}$.

This is equivalent to the statement of the theorem since by setting $\varepsilon = c/(2 + 2\delta)$, we get that there is no efficient $(n^{1+\delta}/(n^{1+\delta} + 1))(1 + n^{-c}) < (1 + n^{-c})$-approximation algorithm for NANIP.

Let G be a graph on $n = 2^{(1+\delta)R}$ vertices containing a k-clique where $k = n^{1/(1+\delta)} = 2^R$ and construct G' from G by adding a k-independent set as before, with $f(i) = \max(k - i, 0)$. Suppose we have an efficient C-approximation algorithm for NANIP. After running it on input (G', k), modify the output sequence according to the previous lemma. Then all the nodes after the first k are free, since after the first k vertices the vertices of U will follow, which are all connected to the first k vertices and are therefore free, and after them every vertex will have at least k neighbours (the k vertices in U) and hence will be free too. Thus the cost of the sequence is determined by the first k vertices. Since they all have fewer than k preceding neighbours, the cost function for them is linear, implying that the total cost of the sequence depends only on the number of edges in between the first k vertices.

The cost of the optimal NANIP sequence in G' is $k(k + 1)/2$, thus the cost of the sequence returned by the approximation algorithm is at most

$$\frac{k}{k+1}\left(1 + \frac{1}{k^{2\varepsilon}}\right) \cdot \frac{k(k+1)}{2} = \frac{1}{2}(k^2 + k^{2-2\varepsilon}).$$

Since

$$\frac{1}{2}(k^2 + k^{2-2\varepsilon}) = (-1)(1 - k^{-2\varepsilon})\frac{k^2}{2} + k^2,$$

it follows by [1], Corollary 2, that there are more than $(1 - k^{-2\varepsilon})k^2/2$ edges between the first k vertices.

Turán's theorem [15] states that, a graph on k vertices that does not contain an $(r + 1)$-clique can have at most $(1 - 1/r)k^2/2$ edges. The contrapositive implies that the induced subgraph on the first k vertices of the NANIP sequence contains a $(k^{2\varepsilon} - 1)$-clique. Since $k^{2\varepsilon} - 1 > 2^{\delta R}$, this means that in a graph that contains a 2^R clique, we can use the C-approximation algorithm for the NANIP instance constructed from the graph to find a clique that is larger than $2^{\delta R}$, thus distinguishing between a clique number of at least 2^R from a clique number of at most $2^{\delta R}$. Thus we reduced this NP-hard problem to C-approximating NANIP, which proves the NP-hardness of the latter problem. □

4. Greedy analysis for convex NANIP

In this section, we discuss the approximation guarantees of the greedy algorithm on convex NANIP. The greedy algorithm is defined to choose the cheapest cost vertex at every step, breaking ties arbitrarily. If the cost function is decreasing, then after an arbitrarily chosen first vertex, the cheapest cost vertex will always have an already installed neighbour. Therefore, the greedy algorithm always produces a connected traversal of a connected graph, in the sense that every prefix of the final traversal induces

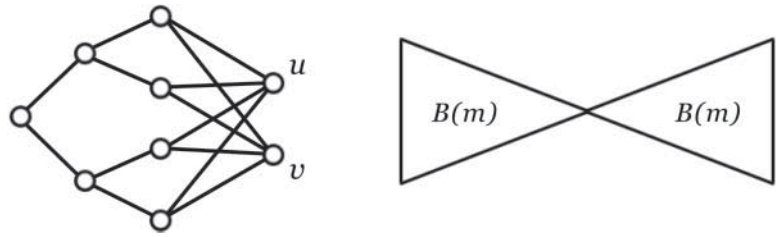


FIG. 2. Left: the graph $B(3)$; right: two $B(m)$ pieced together to force a connected algorithm to incur $\Omega(\log(n))$ cost.

a connected subgraph. We call an algorithm which always produces a connected traversal a *connected algorithm*.

Previous studies of related problems in optimal networks suggested that there is a transition in the structure of the solution as the cost function is made convex. Indeed, it is easy to see that under a non-convex decreasing cost function it is often optimal to use non-connected solutions.

In the convex case, many small instances have optimal solutions that are connected. Therefore, our next theorem shows a rather surprising result, that optimal recovery sometimes requires disconnected solutions, even on convex cost functions. Connected solutions can perform quite badly, having a cost that is a $\Omega(\log n)$ multiple of the optimum.

THEOREM 9 Connected algorithms have an approximation ratio $\Omega(\log(n))$ for convex NANIP problems.

Proof. We construct a particular instance for which a connected algorithm incurs cost $\Omega(\log(n))$ while the optimal route has constant cost. Define the graph $B(m)$ to be a complete binary tree T with m levels, and a pair of vertices u, v such that the leaves of T and $\{u, v\}$ form the complete bipartite graph $K_{2^{m-1},2}$. As an example, $B(3)$ is given in Fig. 2.

Define the cost function $f(n)$ such that $f(0) = 2, f(1) = 1$ and $f(n) = 0$ for all $n \geqslant 2$. For this cost function, it is clear that the minimum cost of a traversal of $B(m)$ is exactly 4 by first choosing the two vertices of $B(m)$ that are not part of the tree, and then traversing the rest of the tree at zero cost. However, if a connected algorithm were forced to start at the root of the tree, it would incur cost $\Omega(m) = \Omega(\log(n))$ since every vertex would have at most one visited neighbour.

To force such an algorithm into this situation we glue two copies of $B(m)$ together so that their trees share a root. Then any connected ordering must start in one of the two copies, and to visit the other copy it must pass through the root, incurring a total cost of $\Omega(\log(n))$. On the other hand, the optimal traversal has total cost 8. □

Further, the greedy algorithm, which simply chooses the cheapest vertex at each step and breaks ties arbitrarily, gives a $\Theta(n)$ approximation ratio in the worst case. To see this, note that in the construction from the theorem the only way a connected algorithm can achieve the logarithmic lower bound is by travelling directly from the root to the leaves. But by breaking ties arbitrarily, the greedy algorithm may visit every interior node in the tree before reaching the leaves, thus incurring a linear cost overall.

5. IP for NANIP

In this section, we describe a new IP formulation of the NANIP problem by adding in Miller–Tucker–Zemlin-type subtour elimination constraints [16]. An IP, of course, does not give a polynomial time

algorithm, but can be sufficiently fast for some instances of practical interest. We then show that this formulation, experimentally, improves on the previous formulation by [1].

5.1 *A new integer program*

In what follows, we will assume that the cost function f is a continuous convex decreasing function $\mathbb{R}^{\geqslant 0} \to \mathbb{R}^{\geqslant 0}$ rather than one $\mathbb{N} \to \mathbb{R}^{\geqslant 0}$. It is necessary to extend f to a continuous function for the LP relaxation to be well defined. While there are many ways to do so, formulating the IP for a general continuous f encapsulates all of them.

For an undirected graph $G = (V, E)$ on $n = |V|$ vertices, and introduce the arc set A by replacing each undirected edge with two directed arcs. For all $(i, j) \in A$ define variables $e_{ij} \in \{0, 1\}$. The choice $e_{ij} = 1$ has the interpretation that i is traversed before j in a candidate ordering of the vertices, or that one chooses the directed edges (i, j) and discards the other. In order to maintain consistency of the IP we impose the constraint $e_{ij} = 1 - e_{ji}$ for all edges (i, j) with $i < j$. Finally, we wish to enforce that choosing values for the e_{ij} corresponds to defining a partial order on V (i.e., that the subgraph of chosen edges forms a DAG). We use the subtour elimination technique of Miller, Tucker and Zemlin [16] and introduce variables u_i for $i = 1, \ldots, n$ with the constraints

$$\begin{aligned} u_i - u_j + 1 &\leqslant n(1 - e_{ij}) \quad && \forall (i, j) \in A \\ 0 \leqslant u_i &\leqslant n && i = 1, \ldots, n \end{aligned} \tag{3}$$

Thus, if i is visited before j then $u_i \geqslant u_j - 1$. Now denote by $d_i = \sum_{(j,i) \in E} e_{ji}$, which is the number of neighbours of v_i visited before v_i in a candidate ordering of V. The objective function is the convex function $\sum_i f(d_i)$, and putting these together we have the following convex integer programme:

$$\begin{aligned} \min \ & \sum_i f(d_i) \\ \text{s.t. } & d_i = \sum_{(j,i) \in A} e_{ji} \quad i = 1, \ldots, n \\ & e_{ij} = 1 - e_{ji} \quad (i, j) \in A, \ \ i < j \\ & u_i - u_j + 1 \leqslant n(1 - e_{ij}) \quad (i, j) \in A \\ & 0 \leqslant u_i \leqslant n \quad i = 1, \ldots, n \\ & e_{ij} \in \{0, 1\} \quad (i, j) \in A \end{aligned}$$

The integer programme has a natural LP relaxation by replacing the integrality constraints with $0 \leqslant e_{ij} \leqslant 1$. Because f is only evaluated at integer points, it is possible to replace $f(d_i)$ with a real-valued variable bound by a set of linear inequalities, as detailed in [1].

5.2 *Experimental results*

We compared the new IP formulation with the formulation of [1] in the algebraic optimization framework. For the comparison, we constructed random connected graphs by first constructing a random tree and then randomly inserting the desired number of edges. For each graph size and order, we constructed five graphs and reported the average running time of the two algorithms. Simulations were run

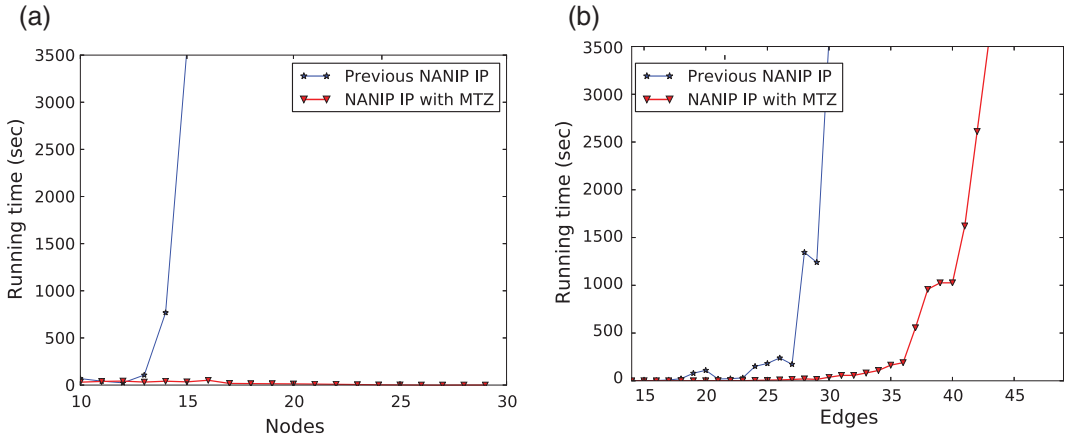


FIG. 3. A comparison of the formulations in [1] and our new IP formulation with MTZ-type constraints. This graph plots running time vs. (a) number of nodes and (b) number of edges in the target graph. In (a) the number of edges was kept at 30 throughout, while in (b) the number of nodes was 15 throughout.

on IBM ILOG CPLEX 12.4 solver running with a single thread on Intel(R) Core(TM) i5 CPU U 520 @ 1.07 GHz with 3.84E6 kB of random access memory.

From the computational experiments it is clear that the MTZ-type formulation gives significant improvements. For instance, the solve time seems to not depend on the number of nodes in the graph (Fig. 3(a)), unlike in the previous formulation. We are also able to solve NANIP instances on 45 edges in under an hour, whereas the previous formulation solved only 30 edge graphs in that span of time (Fig. 3(b)).

6. Conclusion

We analysed the recently introduced NANIP. We proved the NP-hardness of the problem for the practically most relevant case of convex decreasing cost functions, addressing an open problem raised in [1]. We then showed that the worst case approximation ratio of the natural greedy algorithm is $\Theta(n)$. We also gave a new IP formulation for optimally solving NANIP, which outperforms previous formulations.

The approximability of NANIP remains an open problem. In particular, it is still not known whether an efficient $o(n)$ approximation algorithm exists for general convex decreasing cost functions. One obstacle to finding a good rounding algorithm is that the IP we presented has an infinite integrality gap. As proof, the graph K_n with the function $f(i) = \max(0, n/2 - i)$ has $\mathrm{OPT} = \Omega(n^2)$ but the linear relaxation has $\mathrm{OPT}_{\mathrm{LP}} = 0$. So an approximation algorithm via LP rounding would require a different IP formulation.

Funding

A.G. was supported in part by an ORISE fellowship at the Food and Drug Administration. CPLEX software was provided by IBM through the IBM Academic Initiative program.

REFERENCES

1. GUTFRAIND, A., BRADONJIĆ, M. & NOVIKOFF, T. (2015) Modelling the neighbour aid phenomenon for installing costly complex networks. *J. Complex Netw.*, **3**, 249–263.

2. Guha, S., Moss, A., (Seffi) Naor, J. & Schieber, B. (1999) Efficient recovery from power outage (extended abstract). In *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing, STOC'99*. New York, NY: ACM, pp. 574–582.

3. Nurre, S. G. & Sharkey, T. C. (2010) Restoring infrastructure systems: an integrated network design and scheduling problem. In *Proceedings of the 2010 Industrial Engineering Research Conference,* Cancún, Mexico.

4. Lee, E. E., Mitchell, J. E. & Wallace, W. A. (2007) Restoration of services in interdependent infrastructure systems: a network flows approach. *IEEE Trans. Syst. Man Cybernet. Part C: Appl. Rev.*, **37**, 1303–1317.

5. Adibi, M. M. & Fink, L. H. (1994) Power system restoration planning. *IEEE Trans. Syst.*, **9**, 22–28.

6. Bertoli, P., Cimatti, R., Slaney, J. & Thibaux, S. (2002) Solving power supply restoration problems with planning via symbolic model checking. In *AIPS-02 Workshop on Planning via Model-Checking*, pp. 576–580.

7. Coffrin, C., Van Hentenryck, P. & Bent, R. (2011) Strategic stockpiling of power system supplies for disaster recovery. In *Power and Energy Society General Meeting*. San Diego, CA, USA: IEEE, pp. 1–8.

8. Van Hentenryck, P., Bent, R. & Coffrin, C. (2010) Strategic planning for disaster recovery with stochastic last mile distribution. *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems* (A. Lodi, M. Milano & P. Toth eds). Lecture Notes in Computer Science, vol. 6140. Berlin, Heidelberg: Springer, pp. 318–333.

9. Held, M. & Karp, R. M. (1961) A dynamic programming approach to sequencing problems. In *ACM '61: Proceedings of the 1961 16th ACM National Meeting*. New York, NY: ACM, pp. 71.201–71.204.

10. Mitchell, J. E. & Borchers, B. (1996) Solving real-world linear ordering problems using a primal-dual interior point cutting plane method. *Ann. Oper. Res.*, **62**, 253–276.

11. West, D. B. (2001) *Introduction to Graph Theory*. NJ: Pearson Prentice Hall.

12. Schrijver, A. (2005) On the history of combinatorial optimization (till 1960). *Handbooks in Operations Research and Management Science*, vol. 12. Netherlands: Elsevier Science Publ. Co., pp. 1–68.

13. Downey, R. G. & Fellows, M. R. (2013) *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Berlin, Heidelberg, New York: Springer.

14. Zuckerman, D. (2006) Linear degree extractors and the inapproximability of max clique and chromatic number, In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing, STOC'06*. New York, NY: ACM, pp. 681–690.

15. Turán, P. (1941) On an extremal problem in graph theory. *Mat. Fizikai Lapok*, **48**, 436–452.

16. Miller, C. E., Tucker, A. W. & Zemlin, R. A. (1960) Integer programming formulation of traveling salesman problems. *JACM*, **7**, 326–329.