|  |  |
| --- | --- |
| UIC_SPH.PNG  February 20, 2015 |  |

Professor Cris Moore

Santa Fe Institute,

Santa Fe, NM, USA

Dear Dr. Moore:

Thank you very much for editing our prior submission. We are grateful for your insightful editorial suggestions and the careful and constructive evaluations from the anonymous referees.

Below we detail our revisions made in response to the review. In particular, we address each of the comments and articulated more clearly the novelty in our paper as compared to the earlier paper in J Complex Networks.

Our manuscript has not been submitted or accepted for publication at any other journal.

All authors fulfill the criteria. All authors have contributed to, seen, and approved the final, submitted version of the manuscript. No writing assistance other than copy editing was provided in the preparation of the manuscript.

Thank you very much for your consideration.

Sincerely,

Alexander Gutfraind

Jeremy Kun

Adam D. Lelkes

and Lev Reyzin

*University of Illinois at Chicago*

Editor’s comments:

**C**1 In particular, please respond to reviewer #2 and address to what extent this paper improves over ref. [6] (which shares the first author with this submission) and why these results are not a merely incremental improvement.

---------------------- RESPONSE ----------------------

While we are grateful for the volunteer work of Reviewer 2, we believe that they misunderstand our hardness results.

Reviewer 2’s claims that

1. “The strong NP-completeness result comes directly from the max-clique problem.”
2. “One can just state that the max-clique problem is an instance for this problem and move on.”
3. “Along the same lines, any bound for the clique problem would apply to the recovery problem as a lower bound”

Neither statement is factually true.

1. Our construction does show that solving our problem would solve the MAX-CLIQUE problem, but the construction is non-trivial and does not preserve even the exact clique number of the original graph. Unlike with MAX-CLIQUE, it is easy in practice to find optimal and near-optimal solutions of NANIP, as we showed in our simulations.
2. It is easy to see that MAX-CLIQUE is *not* an instance of our problem as claimed. Rather, the nature of the cost function f has a decisive effect on the hardness of NANIP. For example, in the case where the cost function has the convex form f(0)=1 and f(k) = 0 for all k >=1, NANIP reduces to the problem of finding all the connected components, rather than MAX-CLIQUE: once you install any vertex in the connected component of the clique, the rest of the component is installed without additional cost, no matter how hard it is to find the clique in the graph.
3. Moreover, the approximation lower bounds for MAX-CLIQUE that reviewer 2 claims transfer to our setting explicitly *do not* for our construction. In fact, we would find it very surprising if the clique n(1-ε) hardness result held for NANIP.

Reviewer 2 also does not discuss our connectivity result, which shows, contrary to intuition in small graphs, that even when f is decreasing convex, the optimal solutions would not always be connected. This is shown by a counterexample. We feel that finding this counterexample is one of the main contributions of our paper. This counterexample speaks more generally to the nature of convexity in optimal networks, cf. extensive literature below (although these models use a different objective function):

Steffen Bohn and Marcelo O. Magnasco, “Structure, Scaling, and Phase Transition in the Optimal Transport Network”, Phys. Rev. Lett. 98, 088702, February 2007

Eleni Katifori, Gergely J. Szöllősi, and Marcelo O. Magnasco “Damage and Fluctuations Induce Loops in Optimal Transport Networks”, Phys. Rev. Lett. 104, 048704, January 2010

We amplified this point in the revised paper. Finally, we note that the present paper was written by a largely different research team, as compared to [6].

For reviewers 2 and 3:

We agree that the MTZ formulation is not novel, and we have de-emphasized it in our revised write-up. In particular, we changed the title of the paper away from emphasizing the “faster exact formulations” to the title: “Network installation under convex costs”.

The formulation in [6] did not allow us to introduce these types of constraints, and the MTZ result was not a straightforward modification of the IP in [6], but rather based on a new IP construction.

Finally:

We apologize for mis-citing [6]. Our intent was not to disguise the common author in the two papers, but rather, we mistakenly followed the convention of the theoretical computer science community, where authors are cited alphabetically.

**Detailed response**

Reviewers' Comments:

Reviewer #1:

**C**2The illustration in Figure 1 is confusing. Can you please replace it with a better formulation to compute the cost function? The example in [6] is clear.

---------------------- RESPONSE ----------------------

We fixed a typo in the caption that made the definition of f unclear, made the notation clearer, and gave an example for an optimal traversal for (a) to make the figure easier to understand.

**C**3

It is somewhat unusual to refer a paper [6] with its second author (Bradonjic et al.). Please use the first author's name.

---------------------- RESPONSE ----------------------

We fixed the citations. The mistake comes from the convention in theoretical computer science to list the authors alphabetically.

Reviewer: 2

Comments to the Author

**C**4This paper studies the network recovery problem, where the goal is to reconstruct a network one node at a time and the cost of recovery for a node depends on the existing edges on the network. The paper defines the problem shows that it is strongly NP-hard; gives a lower bound for approximation, and proposes a new integer programming formulation that is faster than a previous one.

The paper improves on an existing paper (cited as [6]) published on the same journal in 2014, and the first authors for the two papers are the same. It was puzzling to me why the authors chose to refer to [6] with the name of the second author as Brandojic et al.

---------------------- RESPONSE ----------------------

We explained above that the mistake comes from the convention in theoretical computer science to list the authors alphabetically.

**C**5I find the contributions of this paper to be very incremental and I do not think they are significant enough for a new publication. The strong NP-completeness result comes directly from the max-clique problem. I don't think a separate construction is necessary. One can just state that the max-clique problem is an instance for this problem and move on. Similarly, the lower bound result follows from results on the clique problem. Along the same lines, any bound for the clique problem would apply to the recovery problem as a lower bound. And I do not think stating these would merit a new paper.

---------------------- RESPONSE ----------------------

We respectfully disagree and address this above.

**C**6The "new" integer programming formulation is based on a paper published in 1960, and can be applied to any problem that includes ordering of a larger set ( edges in this case) is dictated by ordering of a smaller set (vertices). These cuts have been applied to other ordering problems such as the traveling salesperson problem. And I would expect any computational study to use such cuts anyway. So I do not think introducing these cuts from 1960 is a significant contribution either.

In summary, I find the contributions of this paper not to be significant enough for publication in this journal.

---------------------- RESPONSE ----------------------

We agree that the MTZ cuts are not novel per se, and de-emphasized them in the revised manuscript.

Reviewer: 3

Comments to the Author

The paper studies an interesting combinatorial formulation for a

network recovery problem developed in an earlier paper. The main

contributions are hardness results for the decision and approximation

versions, under a convex decreasing cost function. They also study

the fixed parameter complexity. An interesting observation is the

structural result about how connectivity changes the cost of

the solution. The final result is a simple IP and its evaluation.

The paper is generally pretty well written.

The technical contributions are borderline. Strengthening them either in terms of the theoretical results or the empirical results would be good.

---------------------- RESPONSE ----------------------

We certainly share the reviewer’s wish to see additional results, but we would argue that this paper resolves many questions surrounding the computational complexity NANIP, which were left open before. In addition, we note that the reviewer missed the non-trivial construction disproving the connectivity conjecture proposed in [6].

**C**7Lemma 1: instead of "as above", it would be better to refer to the proof of theorem 5.

Also, I assume \sigma is the optimal traversal. This should be stated. The

proof needs more clarification. In case 1, why are the u\_i's all free? Case 2

is also not very clear.

---------------------- RESPONSE ----------------------

We replaced “as above” by “as in the proof of Theorem 5.” \sigma does not have to be an optimal traversal. We clarified this in the statement of the lemma. We added more detailed explanations to both cases.

**C**8Proof of theorem 7: it will be good to clarify why the nodes after k are free

if we have a C-approximation. The last part of this proof should also be

explained further.

---------------------- RESPONSE ----------------------

We added the requested clarification.

**C**9section 4, para 1: there might be some functions for which one can construct greedy traversals which are not connected. I believe the correct statement is that there exists a greedy traversal that is connected. This should be clarified.

---------------------- RESPONSE ----------------------

If the cost function is decreasing, all greedy traversals are connected. We added a short explanation of this fact to the paper.

**C**10The experimental results section is pretty sparse. What kinds of graphs were used? The only observation here is about the running time. It might be good to add some more experimental results, and see what the structure of the solution is.

---------------------- RESPONSE ----------------------

We revised our explanation and added detailed description of the graphs being used. From our observations, the structure of the solutions is generally uninteresting: they are in general connected and can often be found by greedy techniques. Given the overall thrust of our paper and the previous results in [6], we felt that this section should be kept short.

**C**11introduction, para 2: Bradonjic et al. should be Gutfraind et al.

- page 3, para 2: "NANIP model" --> "NANIP problem"

- page 3, para 3: "Similarly to ..." --> "Similar to"

---------------------- RESPONSE ---------------------

These mistakes have been fixed-