Pseudo-Separation for Assessment of Structural Vulnerability of a Network

Alan Kuhnle
University of Florida
kuhnle@ufl.edu

Tianyi Pan
University of Florida
tianypan@cise.ufl.edu

Victoria G. Crawford
University of Florida
vcrawford01@ufl.edu

Md Abdul Alim
University of Florida
alim@cise.ufl.edu

My T. Thai
University of Florida
mythai@ufl.edu

ABSTRACT

Based upon the idea that network functionality is impaired if two nodes in a network are sufficiently separated in terms of a given metric, we introduce two combinatorial *pseudocut* problems generalizing the classical min-cut and multi-cut problems. We expect the pseudocut problems will find broad relevance to the study of network reliability. We comprehensively analyze the computational complexity of the pseudocut problems and provide three approximation algorithms for these problems.

Motivated by applications in communication networks with strict Quality-of-Service (QoS) requirements, we demonstrate the utility of the pseudocut problems by proposing a targeted vulnerability assessment for the structure of communication networks using QoS metrics; we perform experimental evaluations of our proposed approximation algorithms in this context.

KEYWORDS

Approximation algorithms; Network cutting problems

ACM Reference format:

Alan Kuhnle, Tianyi Pan, Victoria G. Crawford, Md Abdul Alim, and My T. Thai. 2017. Pseudo-Separation for Assessment of Structural Vulnerability of a Network. In *Proceedings of SIGMETRICS '17, Urbana-Champaign, IL, USA, June 5-9, 2017*, 2 pages.
<https://doi.org/http://dx.doi.org/10.1145/3078505.3078538>

1 INTRODUCTION

Many studies of network vulnerability, or the degree to which the functionality of a network may be disrupted by failures, have incorporated connectivity as a fundamental measure of network functionality [1, 2]. Whatever functionality a network may provide to a pair of nodes is usually absent if the pair is disconnected. In this work, we generalize the idea of network impairment resulting from disconnection of network elements to impairment upon sufficient separation in the network according to a given metric; that is, if $d(s, t) > T$ for two vertices s, t , distance metric d , and threshold T . As an example of the impairment of network functionality under

sufficient separation, consider the Industrial Internet of Things (IIoT), which allows extensive monitoring and control of production facilities in real time [3]. However, the Quality-of-Service (QoS) requirements for control of production systems are very strict [4]. Thus, if two nodes in the IIoT network are sufficiently separated in terms of the QoS metric, the network functionality is impaired.

Contributions: Based upon this idea of functionality impairment under sufficient separation, we introduce novel T -separation analogues to the min-cut and multi-cut problem [5, 6], T -PCUT and T -MULTI-PCUT, respectively. These problems are formally defined in Section 2. In Section 3, we analyze the computational complexity of these two problems, and in Section 4, we present GEN, an $O(\log n)$ -approximation algorithm, and FEN, a $(T + 1)$ -approximation algorithm. In addition, we provide GEST, an efficient, randomized algorithm with probabilistic performance guarantee: with probability $1 - 1/n$, GEST returns a feasible solution with cost within ratio $O(\alpha\delta^T + \log k)$ of optimal, where k is the number of pairs to T -separate, δ is the maximum degree in the graph, and α is user-defined parameter in $(0, 1)$ which also impacts the running time of GEST. An experimental evaluation of these algorithms and a detailed discussion of related work is presented in the full version¹.

2 PROBLEM DEFINITIONS

In this section, we introduce the vertex versions of the pseudocut problems; the edge versions are presented in the full version. Let T be an arbitrary but fixed constant throughout this section. The problems will take as input a triple (G, c, d) , where G is a directed graph $G = (V, E)$; $c : V \rightarrow \mathbf{R}^+$ is a cost function on vertices representing the difficulty of removing each node; and $d : E \rightarrow \mathbf{R}^+$ is a length function on edges. For example, $d(e)$ could be a QoS metric such as latency or packet loss on edge e . Although both c and d may be considered weight functions, we use *cost* for c and *length* for d to avoid confusion. The case when $c(v) = 1$ for all vertices is referred to as *uniform cost*, and the case when $d(e) = 1$ for all edges is referred to as *uniform length*. The distance $d(u, v)$ between two vertices is the length of the d -weighted, directed, and shortest path between u and v ; the cost $c(W)$ of set W of a set of vertices is the sum of the costs of individual vertices in W .

PROBLEM 1 (MINIMUM T -PSEUDOCUT (T-PCUT)). *Given triple (G, c, d) and a pair (s, t) of vertices of G , determine a minimum cost set $W \subset V \setminus \{s, t\}$ of vertices such that $d(s, t) > T$ after the removal of W from G .*

¹The full version of this paper is available at <https://arxiv.org/abs/1704.04555>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGMETRICS '17, June 5-9, 2017, Urbana-Champaign, IL, USA
© 2017 Copyright held by the owner/author(s).
ACM ISBN ACM ISBN 978-1-4503-5032-7/17/06.
<https://doi.org/http://dx.doi.org/10.1145/3078505.3078538>

PROBLEM 2 (MINIMUM T -MULTI-PSEUDOCUT (T-MULTI-PCUT)). Given triple (G, c, d) , and a target set of pairs of vertices of G , $S = \{(s_1, t_1), (s_2, t_2), \dots, (s_k, t_k)\}$, determine a minimum cost set W of vertices such that $d(s_i, t_i) > T$ for all i after the removal of W from G .

In the above two formulations, we emphasize that the threshold T is a fixed constant independent of the input. We will consider simple paths $p = p_0 p_1 \dots p_l \in G$; that is, paths containing no cycles. Let $\mathcal{P}(s_i, t_i)$ denote the set of simple paths p between $(s_i, t_i) \in S$ that satisfy the condition $d(p) \leq T$, and let $\mathcal{P} = \bigcup_{i=1}^k \mathcal{P}(s_i, t_i)$.

3 COMPUTATIONAL COMPLEXITY

In this section, we summarize our results from the full version on the computational complexity of the pseudocut problems. For uniform-length T-PCUT, we provide polynomial-time algorithms for specific special cases, although with arbitrary edge lengths the problem is shown to be NP-hard. For uniform length, cost T-MULTI-PCUT, an approximation-preserving from the vertex cover problem [7] exists.

PROPOSITION 3.1. For $T \leq 3$, T-PCUT with uniform lengths and costs is solvable in polynomial time.

PROPOSITION 3.2. Let D be a constant, T-PCUT($G, (s, t)$) be an instance of T-PCUT for some constant T with uniform lengths and uniform costs. If the maximum degree δ in G satisfies $\delta \leq D$, then the optimal solution W is computable in polynomial time.

THEOREM 3.3. Consider the decision version of 1-PCUT with uniform costs and arbitrary lengths; that is, given problem instance 1-PCUT($G, (s, t)$) with uniform costs and arbitrary lengths, and given constant $D > 0$, determine if a solution $W \subset V$ exists with $|W| \leq D$. This problem is NP-complete.

THEOREM 3.4. Unless $P = NP$, there is no polynomial-time approximation to uniform length, cost T-MULTI-PCUT within a factor of 1.3606, for $T \geq 1$.

4 APPROXIMATION ALGORITHMS

In this section, we present three approximation algorithms for arbitrary vertex cost T-MULTI-PCUT, when the length function on edges is bounded below: $d(e) > q_{min}$ for some constant $q_{min} > 0$. Let $T_0 = T/q_{min}$. Then, all paths in $\mathcal{P}$ can be enumerated in $O(n^{T_0})$. Furthermore, an optimal solution S of vertices must intersect every path in $\mathcal{P}$. Thus, T-MULTI-PCUT may be considered as a special case of the set covering problem [7], where the paths correspond to the elements to be covered and vertices correspond to sets. More details are presented in the full version.

Then, we have two approximation algorithms for set cover that apply: the greedy algorithm (GEN), choosing at each iteration the vertex covering the most paths in $\mathcal{P}$, and the frequency rounding algorithm (FEN) based upon the optimal solution to the linear program corresponding to the set cover instance [7]. The following theorems are proved in the full version.

THEOREM 4.1. GEN achieves a performance guarantee of $O(\log n)$ with respect to the optimal solution with running time bounded by $O(kn^{T_0})$. Furthermore, for each n , there exists an instance of the single

pair PCUT problem where GEN returns a solution of cost greater than a factor $\Omega(\log n)$ of the optimal.

THEOREM 4.2. FEN achieves a performance guarantee of $T_0 + 1$ with respect to the optimal solution.

Although GEN, FEN both run in polynomial time, the enumeration of all paths at most length T may be very expensive. Therefore, we formulate a probabilistic approximation algorithm, GEST, which is similar to GEN except that GEST estimates which node intersects the most paths; to perform this estimation, GEST samples paths in the following way: for each pair $(s, t) \in S$, a path beginning from s is sampled by randomly walking from s until the path exceeds the length T or until t is reached; if t is reached, the sampled path is *valid*. The number $\tau(S)$ of such paths a set S of vertices intersects is estimated using the probability of the path, which is computed iteratively: each step i in the random walk multiplies the probability of the path by $1/n_i$, where n_i is the number of available vertices at step i . If set S intersects a valid sampled path p , with $Pr(p) = \eta$, then the estimated number of paths between (s, t) that S intersects is increased by $1/\eta$. In the full version, we show that this estimator is unbiased and prove the following bound on the number of samples required to get a good estimate $\sigma(S)$.

LEMMA 4.3. Let the number of paths sampled for each $(s, t) \in S$ be at least $L = 3k^2 \log(2n^2)/2\alpha^2$, where $\alpha \in (0, 1)$. Then, given a set $S \subseteq V$ and δ as the maximum degree in G , the inequality $|\tau(S) - \sigma(S)| < \alpha\delta^{T_0}$ holds with probability at least $1 - 1/n^3$.

Lemma 4.3 allows us to prove the following probabilistic performance guarantee for GEST. The proof is detailed in the full version.

THEOREM 4.4. Given an instance (G, c, d, S) of uniform vertex cost T-MULTI-PCUT whose length function d is bounded below, let δ be the maximum degree in G , and let $\alpha \in (0, 1)$. With probability at least $1 - 1/n$, GEST returns a feasible solution W with cost within ratio $O(\alpha\delta^{T_0} + \log |S|)$ of optimal. The running time of GEST is $O(k^3 n \log(2n^2)/2\alpha^2)$.

5 ACKNOWLEDGEMENTS

This work is supported in part by NSF grant CNS-1443905 and DTRA grant HDTRA1-14-1-0055.

REFERENCES

- [1] Tony H Grubbs, Timothy C Matisziw, Alan T Murray, and Diane Snediker. Comparative Approaches for Assessing Network Vulnerability. *International Regional Science Review*, 31(1):88–112, 2008.
- [2] Thang N. Dinh and My T. Thai. Network under joint node and link attacks: Vulnerability assessment methods and analysis. *IEEE/ACM Transactions on Networking*, 23(3):1001–1011, 2015.
- [3] Ahmad-Reza Sadeghi, Christian Wachsmann, and Michael Waidner. Security and Privacy Challenges in Industrial Internet of Things. *Proceedings of the 52nd Annual Design Automation Conference - DAC '15*, 17:1–6, 2015.
- [4] Linus Thrybom and Gunnar Prytz. QoS in Switched Industrial Ethernet. *IEEE Conference on Emerging Technologies and Factory Automation*, 2009.
- [5] Naveen Garg, Vijay V. Vazirani, and Mihalis Yannakakis. Multiway Cuts in Directed and Node Weighted Graphs. In *International Colloquium on Automata, Languages, and Programming*, Berlin, 1994. Springer-Verlag.
- [6] Tom Leighton and Satish Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, 46(6):787–832, 1999.
- [7] Vijay V Vazirani. *Approximation Algorithms*. Springer-Verlag Berlin Heidelberg, first edition, 2003.