2B-1

On the complexity of Minimum Topic-Connected Overlay Problems

Juraj Hromkovič[‡] Taisuke Izumi[†] Hirotaka Ono[⋆] Monika Steinová[‡] Koichi Wada[†*]
† Graduate School of Engineering, Nagoya Institute of Technology, Japan
‡ Department of Computer Science, ETH Zurich, Switzerland
⋆ Department of Economic Engineering, Kyushu University, Japan

1 Introduction

In the context of designing a scalable overlay network to support decentralized topic-based pub/sub communication, the Minimum Topic-Connected Overlay problem (Min-TCO in short) has been investigated[2, 3]: Given a set of t topics and a collection of n users together with the lists of topics they are interested in, the aim is to connect these users to a network by a minimum number of edges such that every graph induced by users interested in a common topic is connected. It is known that Min-TCO is $\mathcal{NP}$-hard and approximable within $O(\log t)$ in polynomial time[2, 1].

In this paper, we further investigate the problem and some of its special instances. We give various hardness results for instances where the number of topics in which an user is interested in is bounded by a constant, and also for the instances where the number of users interested in a common topic is constant. For the latter case, we present a first constant approximation algorithm. We also present some polynomial-time algorithms for very restricted instances of Min-TCO.

2 Minimum Topic-Connected Overlay Problem

The set of users or nodes of our network is denoted by $U = \{u_1, u_2, \ldots, u_n\}$. The topics are $T = \{t_1, t_2, \ldots, t_m\}$. Each user subscribes to several topics. This relation is expressed by the user interest function $\mathrm{INT} : U \to 2^T$. The set of all vertices of U interested in a topic t is denoted by U_t. For a given set of users U, a set of topics T, and an interest function INT, we say that a graph $G = (U, E)$ with $E \subseteq \{\{u, v\} \mid u, v \in U \wedge u \neq v\}$ is *t-topic-connected*, for $t \in T$, if the subgraph $G[U_t]$ is connected. We call the graph *topic-connected* if it is t-topic-connected for each topic $t \in T$. Note that the topic-connectedness property implies that a message published for topic t is transmitted to all users interested in this topic without using non-interested users as intermediate nodes.

The most general problem that we study in this paper is called *Minimum Topic Connected Overlay*:

Problem 1 Min-TCO *is the following optimization problem:*

Input: *A set of users U, a set of topics T, and an user interest function* $\mathrm{INT} : U \to 2^T$.

Feasible solutions: *Any set of edges $E \subseteq \{\{u, v\} \mid u, v \in U \wedge u \neq v\}$ such that the graph (U, E) is topic-connected.*

Costs: *Size of E.*

Goal: *Minimization.*

We refer here to the famous *minimum hitting set problem* (Min-HS).In Min-HS, we are given a system of sets $\mathcal{S} = \{S_1, \ldots, S_m\}$ on n elements $X = \{x_1, \ldots, x_n\}$ (i. e., $S_j \subseteq X$). A feasible solution of this problem is a set $H \subseteq X$, such that $S_j \cap H \neq \emptyset$ for all j. Our goal is to minimize the size of H. In our paper, we refer to the d-HS problem – a restriction of Min-HS to instances where $|S_i| \leq d$ for all i.

*email:wada@nitech.ac.jp

3 Results for Min-TCO When $|U_t|$ is Constant

3.1 Hardness results

Theorem 1 *If* $\max_{t \in T} |U_t| \le 2$*, then* Min-TCO *can be solved in linear time.*

Theorem 2 *For arbitrary* $d \ge 2$*, there exists an AP-reduction from* d-HS *to* Min-TCO*, where* $\max_{t \in T} |U_t| \le d+1$.

Corollary 1 *For any* $\delta > 0$ *and polynomial-time* α*-approximation algorithm of* Min-TCO *with* $\max_{t \in T} |U_t| \le d+1$*, there exists a polynomial-time* $(\alpha+\delta)$*-approximation algorithm of* d-HS.

Corollary 2 Min-TCO *with* $\max_{t \in T} |U_t| \le d$ *(*$d \ge 3$*) is* $\mathcal{NP}$*-hard to approximate within a factor of* $(d-1-\varepsilon)$*, for any* $\varepsilon > 0$*, and, if the unique games conjecture holds, there is no polynomial-time* $(d-\varepsilon)$*-approximation algorithm for it.*

Corollary 3 Min-TCO *is* $\mathcal{LOGAPX}$*-complete.*

3.2 A Constant Approximation Algorithm

Theorem 3 *There exists a one-to-one reduction of instances of* Min-TCO *with* $\max_{t \in T} |U_t| \le d$ *to instance of* $O(d^2)$-HS.

Theorem 4 *There exists a polynomial-time* $(\lfloor d/2 \rfloor \cdot \lceil d/2 \rceil)$*-approximation algorithm for* Min-TCO *with* $\max_{t \in T} |U_t| \le d$.

Corollary 4 Min-TCO *with* $\max_{t \in T} |U_t| \le 3$ *inherits the approximation hardness of* 2-HS[1].

Corollary 5 Min-TCO *with* $\max_{t \in T} |U_t| \le d$ *is* $\mathcal{APX}$*-complete, for arbitrary* $d \ge 3$.

4 Hardness of Min-TCO When $|INT(v)|$ is Constant

Theorem 5 Min-TCO *with* $\max_{v \in U} |\mathrm{INT}(v)| \le 6$ *cannot be approximated within a factor of* 694/693 *in polynomial time, unless* $\mathcal{P} = \mathcal{NP}$*, even if* $|\mathrm{INT}(v) \cap \mathrm{INT}(u)| \le 3$ *holds for every pair of different users* $u, v \in U$.

[1] 2-HS is equevalent to the minimum vertex cover problem.

Corollary 6 Min-TCO *with* $\max_{v \in U} |\mathrm{INT}(v)| \le 6$ *is* $\mathcal{APX}$*-complete.*

Corollary 7 Min-TCO *with* $|\mathrm{INT}(v) \cap \mathrm{INT}(u)| \le 3$*, for all users* $u, v \in U$*, is* $\mathcal{APX}$*-complete.*

Theorem 6 Min-TCO *can be solved in linear time, if* $|\mathrm{INT}(v) \cap \mathrm{INT}(u)| \le 2$ *holds for every pair of users* $u, v \in U$*,* $u \ne v$.

Corollary 8 Min-TCO *with* $\max_{u \in U} |\mathrm{INT}(u)| \le 3$ *can be solved in linear time.*

5 A Polynomial-Time Algorithm for Min-TCO with Bounded Number of Topics

Theorem 7 *The optimal solution of* Min-TCO *can be computed in polynomial time if* $|T| \le (1+\varepsilon(|U|))^{-1} \cdot \log\log |U|$*, for a function*

$$\varepsilon(n) \ge \frac{3/2 \log\log\log n}{\log\log n - 3/2 \log\log\log n}.$$

References

[1] Dana Angluin, James Aspnes, and Lev Reyzin. Inferring social networks from outbreaks. In *Proc. of the 21st International Conference on Algorithmic Learning Theory (ALT 2010)*, volume 6331 of *Lecture Notes in Computer Science*, pages 104–118. Springer-Verlag, 2010.

[2] Gregory Chockler, Roie Melamed, Yoav Tock, and Roman Vitenberg. Constructing scalable overlays for pub-sub with many topics. In *Proc. of the 26th Annual ACM Symposium on Principles of Distributed Computing (PODC 2007)*, pages 109–118. ACM, 2007.

[3] Melih Onus and Andréa W. Richa. Minimum maximum degree publish-subscribe overlay network design. In *Proc. of IEEE INFOCOM 2009*, pages 882–890. IEEE, 2009.