A Game Theoretic Approach For k-Core Minimization

Extended Abstract

Sourav Medya
Northwestern University
sourav.medya@kellogg.northwestern.edu

Tianyi Ma
University of California Los Angeles
tianyima@cs.ucla.edu

Arlei Silva
University of California Santa Barbara
arlei@cs.ucsb.edu

Ambuj Singh
University of California Santa Barbara
ambuj@cs.ucsb.edu

ABSTRACT

K-cores are maximal induced subgraphs where all vertices have degree at least k. These dense patterns have applications in community detection, network visualization and protein function prediction. However, k-cores can be quite unstable to network modifications, which inspires the question: *How resilient is the k-core structure of a network, such as the Web or Facebook, to edge deletions?* More specifically, we study *the problem of computing a small set of edges for which the removal minimizes the k-core structure of a network.* This paper provides a comprehensive characterization of the hardness of the k-core minimization problem (KCM), including innaproximability and parameterized complexity. Motivated by these challenges, we propose a novel algorithm inspired by Shapley value—a cooperative game-theoretic concept— that is able to leverage the strong interdependencies in the effects of edge removals in the search space. Our experiments, show that the proposed algorithm outperforms competing solutions in terms of k-core minimization.

KEYWORDS

k-core, network design, network resilience, shapley value

ACM Reference Format:
Sourav Medya, Tianyi Ma, Arlei Silva, and Ambuj Singh. 2020. A Game Theoretic Approach For k-Core Minimization. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

K-cores play an important role in revealing the higher-order organization of networks. A k-core [10] is a maximal induced subgraph where all vertices have internal degree of at least k. These cohesive subgraphs have been applied to model users' engagement and viral marketing in social networks [2]. Other applications include anomaly detection [12], community discovery [8], and visualization [3]. However, the k-core structure can be quite unstable under network modification. For instance, removing only a few edges might lead to the collapse of the core structure of a graph. This motivates the k-core minimization problem: *Given a graph G and constant k, find a small set of b edges for which the removal minimizes the size of the k-core structure [15].*

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

We motivate k-core minimization using the following applications: (1) *Monitoring:* Given an infrastructure or technological network, which edges should be monitored for attacks [6]? (2) *Defense:* Which communication channels should be blocked in a terrorist network in order to destabilize its activities [9]? and (3) *Design:* How to prevent unraveling in a social or biological network by strengthening connections between pairs of nodes [2]?

There is no polynomial time algorithm that achieves a constant-factor approximation for KCM. This behavior differs from more popular problems in graph combinatorial optimization, such as submodular optimization, where a simple greedy algorithm provides constant-factor approximation guarantees. The algorithm proposed in this paper applies the concept of *Shapley values* (SVs), which, in the context of cooperative game theory, measure the contribution of players in coalitions [11]. Our algorithm selects edges with largest Shapley value to account for the joint effect (or cooperation) of multiple edges in the solution set.

Recent papers have introduced the KCM problem [15] and its vertex version [13], where the goal is to delete a few vertices such that the k-core structure is minimized. However, our work provides a stronger theoretical analysis and more effective algorithms that can be applied to both problems. In particular, we show that our algorithm outperforms the greedy approach proposed recently in [15]. Our main contributions are summarized as follows:

- We study the k-core minimization (KCM) problem, which consists of finding a small set of edges, removal of which minimizes the size of the k-core structure of a network.
- We show that KCM is NP-hard, even to approximate by a constant for $k \geq 3$. We also discuss the parameterized complexity of KCM and show the problem is $W[2]$-hard parameterized by budget.
- Given the above inapproximability result, we propose a Shapley Value based algorithm that efficiently accounts for the interdependence among the candidate edges for removal. We show the accuracy and efficiency of our algorithm using several datasets.

Due to space limitation we defer some details, proofs, and experimental results to an extended version [7].

Related Work : Adiga et al. [1] studied the stability of high cores in noisy networks. A few works [6, 14] recently introduced a notion of resilience in terms of the stability of k-cores against deletion of random nodes/edges. Another related paper [13] studied the node version of KCM. Bhawalkar et al. [2] and Chitnis et al. [4] studied the problem of increasing the size of k-core by anchoring a few vertices initially outside of the k-core.

2 PROBLEM DEFINITION

We assume $G(V, E)$ to be an undirected and unweighted graph with sets of vertices V ($|V| = n$) and edges E ($|E| = m$). Let $d(G, u)$ denote the degree of vertex u in G. An induced subgraph, $H = (V_H, E_H) \subset G$ is such that if $u, v \in V_H$ and $(u, v) \in E$ then $(u, v) \in E_H$. The k-core [10] of a network is defined below.

Definition 2.1. k**-Core:** The k-core of a graph G, denoted by $C_k(G) = (V_k(G), E_k(G))$, is defined as a maximal induced subgraph that has vertices with degree at least k.

Let $G^B = (V, E \setminus B)$ be the modified graph after deleting a set B with b edges. Deleting an edge reduces the degree of two vertices and possibly their core numbers. The reduction in core number might propagate to other vertices. For instance, the vertices in a simple cycle are in the 2-core but deleting any edge moves all the vertices to the 1-core. Let $N_k(G) = |V_k(G)|$ and $M_k(G) = |E_k(G)|$ be the number of nodes and edges respectively in $C_k(G)$.

Definition 2.2. **Reduced** k**-Core:** A reduced k-core, $C_k(G^B)$ is the k-core in G^B, where $G^B = (V, E \setminus B)$.

Definition 2.3. K**-Core Minimization (KCM):** Given a candidate edge set Γ, find set $B \subset \Gamma$ of b edges to be removed such that $C_k(G^B)$ is minimized, or, $f_k(B) = N_k(G) - N_k(G^B)$ is maximized.

Inapproximability: The hardness of the KCM problem stems from the fact that there is a combinatorial number of choices of edges from the candidate set, and there might be strong dependencies in the effects of edge removals. KCM is proved to be NP-hard in [15]. We show a stronger result that KCM is NP-hard to approximate within any constant factor.

THEOREM 1. *The KCM problem is NP-hard to approximate within a constant-factor for all* $k \geq 3$.

Theorem 1 shows that there is no polynomial-time constant-factor approximation for KCM when $k \geq 3$. This contrasts with well-known NP-hard graph combinatorial problems in the literature [5]. We explore the hardness of our problem further in terms of exact exponential algorithms with respect to the parameters and show that KCM is $W[2]$-hard.

THEOREM 2. *The KCM problem is* $W[2]$*-hard parameterized by the budget* b.

3 ALGORITHMS AND EXPERIMENTS

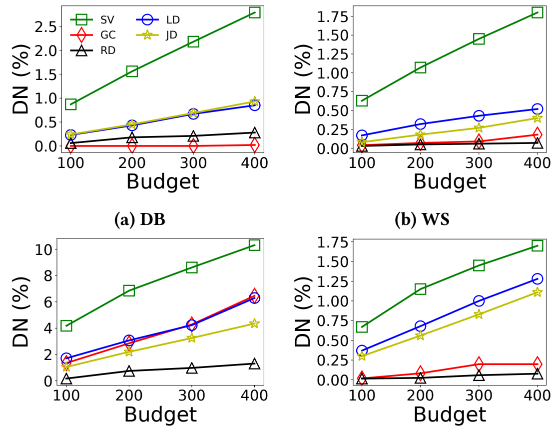
According to the Theorems 1 and 2, an optimal solution— or constant-factor approximation—for k-core minimization requires enumerating all possible size-b subsets from the candidate edge set, assuming $P \neq NP$. Here we propose an efficient heuristic for KCM. The general greedy algorithm is unaware of some dependencies between the candidates in the solution set. To capture these dependencies, we adopt a cooperative game theoretic concept named Shapley Value [11]. Our goal is to make a coalition of edges (players) and divide the total gain by this coalition equally among the edges inside it.

The Shapley Value of an edge e in KCM is defined as follows. Let the value of a coalition P be $\mathcal{V}(P) = f_k(P) = N_k(G) - N_k(G^P)$. Given an edge $e \in \Gamma$ and a subset $P \subseteq \Gamma$ such that $e \notin P$, the marginal contribution of e to P is: $\mathcal{V}(P \cup \{e\}) - \mathcal{V}(P), \quad \forall P \subseteq \Gamma$. Let Ω be the set of all $|\Gamma|!$ permutations of all the edges in Γ and $P_e(\pi)$ be the set of all the edges that appear before e in a permutation π. The Shapley Value of e is the average of its marginal contributions to the edge set that appears before e in all the permutations:

$$\Phi_e = \frac{1}{|\Gamma|!} \sum_{\pi \in \Omega} \mathcal{V}(P_e(\pi) \cup \{e\}) - \mathcal{V}(P_e(\pi)). \quad (1)$$

Shapley Values capture the importance of an edge inside a set (or coalition) of edges. However, computing Shapley Value requires considering $O(|\Gamma|!)$ permutations. We efficiently approximate the Shapley Value for each edge via sampling.

We evaluate the proposed Shapley Value based algorithm (SV) for k-core minimization against baseline solutions. The experiments were conducted on a 2.59 GHz Intel Core i7-4720HQ machine with 16 GB RAM running Windows 10. Algorithms were implemented in Java. The datasets (EE: Enron, DB: DBLP, FB: Facebook, and WS: Web-Stanford) are available online[1].



(a) DB (b) WS (c) EE (d) FB

Figure 1: K-core minimization (DN(%)) varying the number of edges in the budget: The Shapley Value based algorithm (SV) outperforms the best baseline (LD) by up to 6 times.

Besides the Greedy Cut (GC) algorithm [15], we also consider three more baselines in our experiments. *Low Jaccard Coefficient (JD)* removes the k edges with lowest Jaccard coefficient. Similarly, *Low-Degree (LD)* deletes k edges for which adjacent vertices have the lowest degree. We also apply *Random (RD)*, which simply deletes k edges from the candidate set Γ uniformly at random. The evaluation measure is the percentage $DN(\%)$ of vertices from the initial graph G that leave the k-core after the deletion of edges in B: $DN(\%) = \frac{N_k(G) - N_k(G^B)}{N_k(G)} \times 100$.

Figure 1 presents the k-core minimization results for $k = 5$ using four different datasets. SV outperforms the best baseline by up to six times as our algorithm can capture strong dependencies among sets of edges. On the other hand, GC, which takes into account only marginal gains for individual edges, achieves worse results than simple baselines such as JD and LD. A large value of k leads to a less stable k-core that can be broken by the removal of edges with low-degree endpoints. LD is a good alternative for such extreme scenarios.

[1] https://snap.stanford.edu

REFERENCES

[1] Abhijin Adiga and Anil Kumar S Vullikanti. 2013. How robust is the core of a network?. In *ECML/PKDD*. Springer, 541–556.
[2] Kshipra Bhawalkar, Jon Kleinberg, Kevin Lewi, Tim Roughgarden, and Aneesh Sharma. 2015. Preventing unraveling in social networks: the anchored k-core problem. *SIAM Journal on Discrete Mathematics* 29, 3 (2015), 1452–1475.
[3] Shai Carmi, Shlomo Havlin, Scott Kirkpatrick, Yuval Shavitt, and Eran Shir. 2007. A model of Internet topology using k-shell decomposition. *PNAS* (2007).
[4] Rajesh Hemant Chitnis, Fedor V Fomin, and Petr A Golovach. 2013. Preventing Unraveling in Social Networks Gets Harder. In *AAAI*.
[5] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *SIGKDD*. 137–146.
[6] Ricky Laishram, Ahmet Erdem Sariyüce, Tina Eliassi-Rad, Ali Pinar, and Sucheta Soundarajan. 2018. Measuring and Improving the Core Resilience of Networks. In *The Web Conference*.
[7] Sourav Medya, Tiyani Ma, Arlei Silva, and Ambuj Singh. 2019. K-core minimization: A game theoretic approach. *arXiv preprint arXiv:1901.02166* (2019).
[8] Chengbin Peng, Tamara G Kolda, and Ali Pinar. 2014. Accelerating community detection by using k-core subgraphs. *arXiv preprint arXiv:1403.2226* (2014).
[9] Arie Perliger and Ami Pedahzur. 2011. Social network analysis in the study of terrorism and political violence. *PS: Political Science & Politics* 44, 1 (2011).
[10] Stephen B Seidman. 1983. Network structure and minimum degree. *Social networks* 5, 3 (1983), 269–287.
[11] Lloyd S Shapley. 1953. A value for n-person games. *Contributions to the Theory of Games* (1953).
[12] Kijung Shin, Tina Eliassi-Rad, and Christos Faloutsos. 2016. CoreScope: Graph Mining Using k-Core Analysis—Patterns, Anomalies and Algorithms. In *ICDM*. 469–478.
[13] Fan Zhang, Ying Zhang, Lu Qin, Wenjie Zhang, and Xuemin Lin. 2017. Finding Critical Users for Social Network Engagement: The Collapsed k-Core Problem. In *AAAI*. 245–251.
[14] Zhongxin Zhou, Fan Zhang, Xuemin Lin, Wenjie Zhang, and Chen Chen. 2019. K-Core Maximization: An Edge Addition Approach. In *IJCAI*.
[15] Weijie Zhu, Chen Chen, Xiaoyang Wang, and Xuemin Lin. 2018. K-core Minimization: An Edge Manipulation Approach. In *CIKM*. ACM, 1667–1670.