Basic information

Name	**SUN, Yahui**
Email	yahuisun@ruc.edu.cn
Languages	Mandarin, English, C++, R, MATLAB
Current position	Lecturer, Department of Computer Science and Technology, School of Information, Renmin University of China

Education and positions

2021-	**Lecturer, School of Information, Renmin University of China**
2020-2021	**Research fellow, School of Computing, National University of Singapore**
2019-2020	**Research fellow, School of Computer Science and Engineering, Nanyang Technological University**
2018-2019	**Postdoctoral fellow, Research School of Engineering, Australian National University**
2014-2018	**Ph.D., School of Electrical, Mechanical and Infrastructure Engineering, University of Melbourne**
	Thesis title: Classical, prize-collecting and node-weighted Steiner tree problems in graphs
2012-2014	**M.S., Department of Aerospace Engineering, Harbin Institute of Technology**
2008-2012	**B.S., Department of Aerospace Engineering, Harbin Institute of Technology**

Research interests

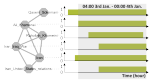
Graph computing and related data mining (such as knowledge graphs and social networks; **I particularly like researches that blur the boundary between applied data engineering and theoretical computer science**)

Career profile

Motivated by China's first crewed space mission in 2003, I studied aerospace engineering for my bachelor and master degrees in the Harbin Institute of Technology. Then, I studied Steiner trees in the field of graph theory for my PhD degree in the University of Melbourne. My research interests originate from my PhD study.

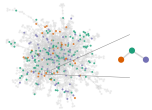
Representative publications

Yahui Sun, Shuai Ma, and Bin Cui. "Hunting temporal bumps in graphs with dynamic vertex properties", Proceedings of the 2022 ACM SIGMOD international conference on management of data (2022).
[**PDF**]



Given a time interval and a graph where vertices exhibit a property of interest (PoI) dynamically, an interesting question is: where (i.e., which part of the graph) and when (i.e., which time sub-interval) does the PoI occur frequently? To our knowledge, no work has been done to answer this question to date. We address this issue in this paper. First, we propose two approximation algorithms. Then, we propose two heuristic algorithms that produce similar solutions with, and are considerably faster than, the two approximation algorithms. Experiments on real datasets show that, in comparison with baselines built using related existing techniques, our algorithms are more suitable for answering the aforementioned question.

Yahui Sun, Xiaokui Xiao, Bin Cui, Saman Halgamuge, Theodoros Lappas, and Jun Luo. "Finding group Steiner trees in graphs with both vertex and edge weights", Proceedings of the VLDB Endowment (2021).
[**PDF**]

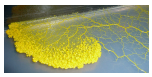


Finding group Steiner trees is a standard approach to information retrieval in relational databases. Most existing work focuses on finding group Steiner trees in vertex-unweighted graphs, and not enough work has been done to find group Steiner trees in graphs with both vertex and edge weights. Here, we develop several algorithms to address this issue. Initially, we extend two algorithms from vertex-unweighted graphs to vertex- and edge-weighted graphs. Then, we develop several new approximation algorithms, one of which provides the tightest polynomial-time approximation guarantee to date. Experiments show that, while no algorithm is the best in all cases, our algorithms considerably outperform the state of the art in many scenarios.

Yahui Sun, Jun Luo, Theodoros Lappas, Xiaokui Xiao, and Bin Cui. "Hunting multiple bumps in graphs", Proceedings of the VLDB Endowment (2020).
[**PDF**]

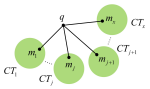
Bump hunting is a graph-related anomaly detection approach. A single bump is hunted in an unweighted graph in the previous work. We extend the previous work by hunting multiple bumps in a weighted graph. We prove that our extended problem can be transformed to a recently formulated prize-collecting Steiner forest problem. We further prove that this problem is NP-hard even in trees. Subsequently, we propose a fast approximation algorithm for solving this problem in trees. Based on this algorithm, we improve the state-of-the-art approximation algorithm for solving this problem in graphs. Experiments on real datasets show the dominance of our improvement over the state-of-the-art algorithms for hunting multiple bumps in graphs.

Yahui Sun, Daniel Rehfeldt, Marcus Brazil, Doreen Thomas, and Saman Halgamuge. “A Physarum-inspired algorithm for minimum-cost relay node placement in wireless sensor networks”, IEEE/ACM Transactions on Networking (2020).
[**PDF**]



Relay node placement is essential in minimizing the costs of wireless sensor networks. Here, we focus on minimum-cost relay node placement. By considering the heterogeneous production and placement costs of relay nodes, our work extends the previous work that considers the costs of relay nodes to be homogeneous. Initially, we conduct some theoretical analyses on the emerging Physarum-inspired algorithms to reveal their potential of computing efficient networks. Based on these analyses, we propose an algorithm for minimum-cost relay node placement. In comparison with the state of the art, our algorithm designs wireless sensor networks with lower relay costs and similar qualities of service. Our work is particularly useful in budget-limited scenarios.

Yahui Sun, Marcus Brazil, Doreen Thomas, and Saman Halgamuge. "The fast heuristic algorithms and post-processing techniques to design large and low-cost communication networks", IEEE/ACM Transactions on Networking (2019).
[**PDF**]



Solving the prize-collecting Steiner tree problem is useful in various scenarios, including computer networking and data mining. We propose two fast algorithms for solving this problem: the first one is a quasilinear-time heuristic algorithm that is faster and consumes less memory than the other algorithms; and the second one is an improvement of a state-of-the-art polynomial-time approximation algorithm (we improve the time complexity of the inside pruning method from $O(n^2)$ to $O(n)$, without sacrificing the optimality of solutions). We show the competitiveness of our algorithms by comparing them with the state of the art. We also propose some post-processing techniques that update the best-known solution for a notoriously difficult benchmark instance.

The other publications

Yahui Sun, and Saman Halgamuge. "Minimum-cost heterogeneous node placement in wireless sensor networks." IEEE Access (2019).
[**PDF**]

Yahui Sun, Chenkai Ma, and Saman Halgamuge. "The node-weighted Steiner tree approach to identify elements of cancer-related signaling pathways." International Conference on Bioinformatics (2017).
[**PDF**]

Yahui Sun, Pathima Nusrath Hameed, Karin Verspoor, and Saman Halgamuge. "A physarum-inspired prize-collecting Steiner tree approach to identify subnetworks for drug repositioning." International Conference on Bioinformatics (2016).
[**PDF**]

Yahui Sun, and Saman Halgamuge. "Fast algorithms inspired by physarum polycephalum for node weighted steiner tree problem with multiple terminals." In IEEE Congress on Evolutionary Computation (2016).
[**PDF**]

Yahui Sun, Yunhai Geng, and Shuang Wang. "Analysis and calibration of star sensor’s image plane displacement." Infrared and Laser Engineering (2014).
[**PDF**]

Yahui Sun, Yingying Xiao, and Yunhai Geng. "On-orbit calibration of star sensor based on a new lens distortion model." In Proceedings of the 32nd Chinese Control Conference (2013).
[**PDF**]

Professional services

Invited reviewer of ACM SIGKDD Conference on Knowledge Discovery and Data Mining, International Conference on Database Systems for Advanced Applications (DASFAA), IEEE Transactions on Communications, IEEE Systems Journal, IEEE Wireless Communications Letters, Artificial Intelligence Review, and Plus One.

(timestamp: 02/2022)