Social network analysis, Material discovery, Urban planning

Efficient and accurate graph data analytics

Multimedia Engineering, Graduate School of Information Science and Technology
Assistant Professor Yuya Sasaki
Researchmap https://researchmap.jp/sasakiyuya?lang=en

Abstract

Graph data is a data structure that represents the states of objects and relationships between objects. We utilize graph data in various applications, such as recommendations on Web shopping and new material discovery on catalysts and crystals (Figure 1). For such diverse and large-scale graphs, I study (1) efficient management and fast search techniques on graph databases, (2) data mining techniques to find new knowledge, and (3) neural networks for accurate prediction (Figure 2).

Background & Results

Relationships between objects can be modeled by graphs, and graphs are widely utilized in our daily life applications. For example, knowledge graphs (structured knowledge of relationships between objects) are utilized in Web searches (e.g., Google search shows tables on the right side when searching for a certain person) and recommendation systems (e.g., related products on Amazon Web shopping). Molecular data represent atoms and the connections between atoms, and we can use them to search for molecules with the same structure. Other applications, such as social network services and road networks, are closely related to our daily lives, and the graphs are becoming larger and more diverse.

I study technology for diverse and large-scale graphs from four points of view: management, search, discovery, and prediction. More concretely, I develop (1) database technology that enables efficient management and fast search on graph data, (2) data mining technology to discover new knowledge, and (3) deep learning technology for accurate predictions (Figure 2).

To my research results, first, we have achieved a speedup of more than 1000 times faster than queries in graph databases by precomputing the structural features of graph data. Second, we developed techniques that can discover exceptional facts and assess discriminatory biases in graph data. Third, we have developed deep learning techniques that can be applied to large-scale graph data with small memory usage and fast learning while keeping high accuracy. Furthermore, in addition to basic development, we are also applying these graph management and analysis techniques to applied fields such as material discovery and medical AI by collaborating with researchers in other fields.

Significance of the research and Future perspective

We are in the Big-data era, and thus technology for efficiently handling large and diverse data is becoming more and more important. Our lives are closely intertwined with various relationships, such as human relationships, jobs, objects owned by people, access to the Web, and agreements between nations. These relationships are converted into graph data. Thus, for understanding and developing the world, it is necessary to predict the state of objects from these relationships and relationships from the state of objects, manage graph data and discover new knowledge from them. I will develop a platform to support easy analysis and management of graph data, and promote the development of technology that can efficiently and accurately analyze large-scale and diverse relationships in the past, present, and future. I contribute to the realization of a more comfortable and happier world.

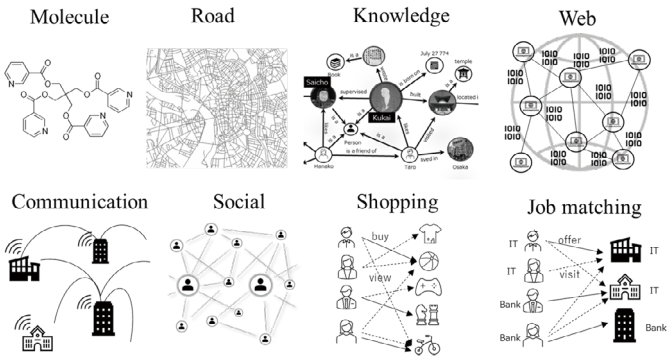

Figure 1. Examples of graph data

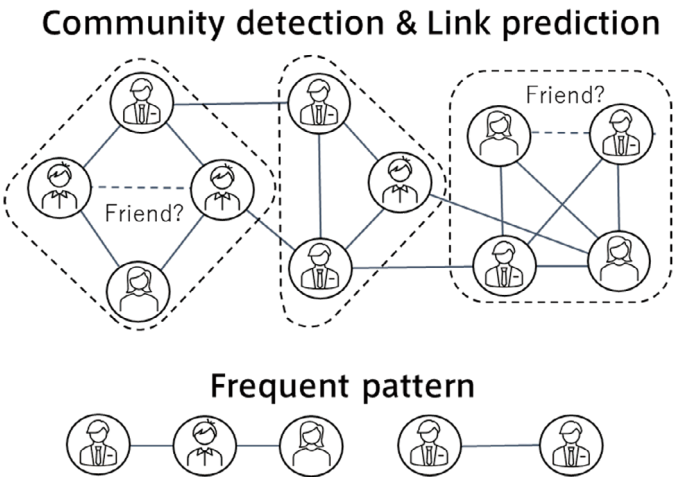

Figure 2. Examples of graph analysis

Patent

Treatise

Maekawa, Seiji; Noda, Koki; Sasaki, Yuya et al. Beyond Real-world Benchmark Datasets: An Empirical Study of Node Classification with GNNs. Conference on Neural Information Processing Systems (NeurIPS), Nov. 2022. doi: 10.48550/arXiv.2206.09144
Sasaki Yuya. Cost-constrained Minimal Steiner Tree Enumeration. Proceedings of International Conference on Information and Knowledge Management (CIKM), pp. 4439-4443, Oct. 2022. doi: 10.1145/3511808.3557570
Sasaki Yuya, George Fletcher, and Onizuka Makoto. Language-aware Indexing for Conjunctive Path Queries. Proceedings of International Conference on Data Engineering (ICDE), pp. 661-673, May 2022. doi: 10.1109/ICDE53745.2022.00054

URL

Keyword graph neural network, big data, data mining, database