Graduate School of Fundamental Science and Engineering
Waseda University

博　士　論　文　概　要
Doctoral Dissertation Synopsis

論　文　題　目
Dissertation Title

A Study of Recommendation Systems with Temporal and Geographical Information

時間的・地理的情報を用いた推薦システムに関する研究

申　請　者
(Applicant Name)
Fan MO
バク　ボン

Department of Computer Science and Communications Engineering, Research on Parallel and Distributed Architecture

November, 2023

1 Introduction:

Recommendation systems return a set of preferable items to a given target user. This thesis aims to improve recommendation accuracy by fully using temporal and geographical information because time and geographical information are common in recommendation systems. Previous works attempted to model geographical information as distance and time information as a sequence. However, they still suffer from insufficient use of side information. This thesis first uses a quantum-inspired computer to model temporal information to improve accuracy by accelerating recommendation optimization and second uses geographical information modeling to assist graph convolution network (GCN) training for higher accuracy. After illustrating the validity of temporal and geographical information, this thesis combines the two to jointly help GCN model training.

2 Objective and Goal:

The research goal of the thesis is to improve recommendation accuracy by modeling temporal and geographical information. The thesis first focuses on shortening model update time to achieve real-time periodic recommendation systems. User preferences change constantly over time. Once the periodic updates can be implemented, the accuracy of estimating user behavior can be improved because the model can capture changes in user behavior in time.

The thesis then works on modeling geographical information to improve recommendation accuracy. Geographical information contains rich user behavioral characteristics. The study is based on the intuition that users in the same geographical area tend to have similar check-in behaviors. A typical application of geographical information is the Point-of-interest (POI) recommendation. Thus, the thesis explores POI recommendations as an example of applied geographical information.

The last goal is to combine temporal and geographical information to co-complement the deep learning models. When combining multiple side of information, it is essential to consider the harmonization between the information. In a graph convolution network, for a target user node, the thesis aims to propose a novel methodology that makes time information to control the graph structure division, i.e., from which nodes to aggregate information while setting geographical information to control the amount of aggregated information, i.e., how much information is aggregated.

3 Challenges:

Challenge 1. Acceleration of Recommendation Optimization

Recommendation optimization - that is, improving accuracy while considering constraints such as budget or cost-per-click (CPC) - is a complex problem in recommendation systems. Achieving high accuracy and quick optimization can be contradictory and present an NP-hard complexity challenge, making periodic updates of the recommendation optimization remain an open question. Although previous works attempted to accelerate optimization, the latency is still too high when dealing with real-time recommendation optimization problems.

Challenge 2. Insufficient Use of Geographical Information

Geographic information, as a latent attribute of POI, holds the potential to improve recommendation accuracy. However, how to fully use geographic information remains an open and challenging question. Previous research ignored that POIs have unique geographic continuity, i.e., multiple POIs can form a geographic area. Simply modeling

the geography as distance in a deep learning model is inadequate, causing loss of area information and thus preventing further enhancement of model performance. When integrating geographic information into deep learning models, mining user behavior over geographic areas can improve the representations of user preferences.

Challenge 3. Insufficient Use of Temporal Information

Time information is generated when a user interacts with an item. Methodologies for modeling temporal information are still in the exploratory phase. In deep learning models, simply modeling time information by sorting user check-ins in chronological order (sequence) cannot fully exploit collaborative signals in time information, which is insufficient. Time information can be divided into time slots to represent the users' preferences during a certain period of time.

3 Contributions:

Contribution 1. Adoption of digital annealers (DA) to accelerate advertisement recommendation and realize periodic recommendation optimization.

Contribution 2. Proposal of active area neighbor to model geographical information in the graph convolution network.

Contribution 3. Proposal of subgraph (time slot) and edge (check-in) propagation-based technique to model time information in the graph convolution network.

Contribution 1. Adoption of digital annealers (DA) to accelerate advertisement recommendation and realize periodic recommendation optimization

Low-speed optimization limits the ability of recommendation models to capture the changes in user interest over time. To solve the problem, the thesis first proposes a periodic recommendation optimization framework. i.e., the model is periodically retained. The proposed real-time recommendation system divides users' behaviors into three stages. 1) Collecting users' data when they visit the websites. 2) Using the collected data to train and optimize the prediction model in the training stage. 3) In the third step, return the optimized result to users and start the next model cycle. Then, a digital annealer (DA) is adopted for acceleration. As a quantum-inspired computer, DA can only support the inputs in a quadratic unconstrained binary optimization (QUBO) model. How to transform a recommendation optimization problem into a QUBO model quickly and thus can be input to the DA to achieve periodic optimization is a challenging task. Specifically, the thesis analyzes the objective function of the recommendation optimization problem, decomposing the objective function to organize the binomial, monomial, and constant terms. These three terms correspond exactly to the binomial, monomial, and constant terms in the QUBO model. Thus, the objective function can be transformed into a QUBA model. Adopting DA with the proposed training technique on advertisement (ad) recommendations results in improved accuracy from 0.3703 to 0.5080 (37.19%) and acceleration of 10.6 times compared to a genetic algorithm-based optimization technique.

Contribution 2. Proposal of active area neighbor to model geographical information in the graph convolution network.

To make full use of geographical information in a lightweight manner, the thesis starts by exploring the users' active areas because POIs are geographically contiguous and thus form areas. To achieve the goal, for each user, the thesis

clusters the POIs that he/she visited to extract the user's active areas. Note that a user may have multiple active areas among cities. The users whose active areas are close (at least one pair of active regions is less than λ km apart) are defined as active area neighbors. Then, the thesis incorporates newly defined neighbors into deep learning, specifically a graph convolution network (GCN), to improve model representation. In a GCN-based recommendation system, a user's neighbors are described as checked items. The thesis extends the traditional definition of neighbor to active area neighbor. As a result, the proposed technique can enhance a GCN model by adopting geographical information, which can extract high-order connectivity over collaborative filtering information. Note that the method does not cause any increase in trainable parameters and keeps the model easy to train. Experiments on the real dataset confirm that the proposed method improves Recall@5 from 0.0788 to 0.0815 on the Gowalla dataset and from 0.0453 to 0.0469 on the Yelp dataset compared with the state-of-the-art LightGCN model.

Contribution 3. Proposal of subgraph (time slot) and edge (check-in) propagation-based technique to model time information in the graph convolution network.

The thesis analyzes and models the time information in the third contribution to train more accurate representations (embeddings) of user preferences. The thesis first divides user check-ins into multiple subgraphs, i.e., time slots, based on time information. In a GCN model, aggregating information only from nodes in the same subgraph enables better mining of users' time-based interests. However, a monotonous subgraph division has drawbacks. i.e., the Monotonous subgraph division cannot propagate the learned time preference features over multiple time slots because the subgraphs are constructed in advance. Thus, the thesis further proposes an edge propagation module to adjust edge affiliation, where edges represent check-ins, to propagate the user's time-based preference to multiple time slots. The propagation module is based on an unsupervised learning algorithm and does not require additional ground-truth labels. This approach to modeling time breaks with the traditional approach of treating temporal information simply as sequence information. Experimental results show that the proposed model further improved Recall@5 to 0.0874 on the Gowalla dataset while from 0.0360 to 0.0388 on the New York dataset compared with state-of-the-art GCN-based models.

4 Organization of the Thesis

In this section, the composition of the thesis is described. The structure of the thesis is listed as follows:

- Chapter 1 introduces the background as well as the goals of the thesis.
- Chapter 2 proposes a real-time periodic advertisement recommendation optimization model using DA. (Contribution 1)
- Chapter 3 describes basic knowledge of POI recommendation, followed by previous works on using time and geographic information.
- Chapter 4 introduces a novel graph convolution network, combining the geographical neighbor concept to model geographical information for POI Recommendation. (Contribution 2)
- Chapter 5 introduces a users' interest propagation-based time-aware graph convolution network to model time information for POI Recommendation. After that, the thesis combines the time aware GCN with geographical information to further improve recommendation performance. (Contribution 3)
- Chapter 6 consists of two parts. The thesis first shows the conclusion of the thesis. Then, the thesis discusses the promising future research directions.

List of research achievements for application of Doctor of Engineering, Waseda University

Full Name : MO FAN seal or signature

Date Submitted(yyyy/mm/dd): 2023/11/02

種類別 (By Type)	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
Journal	○Mo, Fan, and Yamana, Hayato. EPT-GCN: Edge Propagation-Based Time-Aware Graph Convolution Network for POI Recommendation. Neurocomputing, vol. 543, No. 126272, pp.1-15, July 2023, doi:10.1016/j.neucom.2023.126272.
International Conference	○Mo, Fan, and Yamana, Hayato. GN-GCN: Combining Geographical Neighbor Concept with Graph Convolution Network for POI Recommendation. in Proceedings of Information Integration and Web Intelligence: 24th International Conference (iiWAS 2022), pp. 153-165, November 2022, doi:10.1007/978-3-031-21047-1_15. ○Mo, Fan, Jiao, Huida, Morisawa, Shun, Nakamura, Makoto, Kimura, Koichi, Fujisawa, Hisanori, Ohtsuka, Masafumi, and Yamana, Hayato. Real-time Periodic Advertisement Recommendation Optimization under Delivery Constraint using Quantum-inspired Computer. in Proceedings of 2021 International Conference on Enterprise Information Systems (ICEIS 2021), pp. 431-441, April 2021, doi: 10.5220/0010414704310441. Mo, Fan, Jiao, Huida, and Yamana, Hayato. Time Distribution based Diversified Point of Interest Recommendation. in Proceedings of IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA 2020), pp. 37-44, April 2020, doi: 10.1109/ICCCBDA49378.2020.9095741.

List of research achievements for application of Doctor of Engineering, Waseda University

Full Name : MO FAN seal or signature

Date Submitted(yyyy/mm/dd): 2023/11/02

種類別 (By Type)	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む） (theme, journal name, date & year of publication, name of authors inc. yourself)
International Workshop and Poster	Mo, Fan, and Yamana, Hayato. Point of Interest Recommendation by Exploiting Geographical Weigh Center and Categorical Preference. in Proceedings of 2019 International Conference on Data Mining Workshops (ICDMW 2019), pp. 73-76, November 2019, doi: 10.1109/ICDMW.2019.00021. Mo, Fan, Jiao, Huida, Morisawa, Shun, Nakamura, Makoto, Kimura, Koichi, Fujisawa, Hisanori, Ohtsuka, Masafumi, and Yamana, Hayato. Real-Time Periodic Advertisement Recommendation Optimization using Ising Machine. in Proceedings of 2020 IEEE International Conference on Big Data (BigData 2020), pp. 5783-5785, December 2020, doi: 10.1109/BigData50022.2020.9378436.