Mining Business Opportunities from Location-based Social Networks

Shenglin Zhao, Irwin King, Michael R. Lyu*
Department of Computer Science & Engineering
The Chinese University of Hong Kong, Hong Kong
{slzhao,king,lyu}@cse.cuhk.edu.hk

Jia Zeng, Mingxuan Yuan
Huawei Noah's Ark Lab
Hong Kong
{zeng.jia,yuan.mingxuan}@huawei.com

ABSTRACT

Urbanization's rapid progress has modernized a large number of human beings' lives. This urbanization progress is accompanied by the increase of a variety of shops (e.g., restaurants and fitness centers) to meet the increasing citizens, which means business opportunities for the investors. Nevertheless, it is difficult for the investors to catch such opportunities because opening what kind of business at which place is not easy to decide. In this paper, we take this challenge and define the business opportunity mining problem, which recommends new business categories at a partitioned business district. Specifically, we exploit the data from location-based social networks (LBSNs) to mine the business opportunities, guiding the business owners to open new commercial shops in certain categories at a particular area. First, we define the properties of a business district and propose a greedy algorithm to partition a city into different districts. Next, we propose an embedding model to learn latent representations of categories, which captures the functional correlations among business categories. Furthermore, we propose a ranking model based on the pairwise loss to recommend categories for a specific district. Finally, we conduct experiments on Yelp data, and experimental results show that our proposed method outperforms the baseline methods and resolves the problem well.

CCS CONCEPTS

• **Information systems → Location based services**; **Mobile information processing systems**;

KEYWORDS

location-based Services; business intelligence; recommendation system; ranking method

1 INTRODUCTION

Urban development has mostly changed human beings' lives. In the process of urbanization, many people migrate into the cities, and then new business shops are opened in the communities to meet the explosively increasing citizens. Such intensive needs for new shops imply business opportunities for investors. Take the fast food industry as an example, the revenue of this industry has grown from six billion dollars in 1970 to 200 billion dollars in 2015 in the United States[1]. Excluding inflation[2], the revenue in the fast food industry has increased five times, which indicates more consumptions and increasing business shops, implying opportunities for investors.

Nevertheless, discovering business opportunities of new shops is difficult for investors. Opening what kind of business at which place is a tough decision. Even when investors are smart and lucky to find a good location such as a venue near the subway station and try to open a fast food shop for the commuters, it is hard to decide which kind of fast food to provide, sandwiches or fried chickens? The means of surveys can solve this problem to some extent. But the expense is not affordable to most of business owners and investors, except for some big chain retailers.

Fortunately, the emerging location-based social networks (LBSNs) show possibilities to mine business opportunities for investors. LBSNs play important roles in each citizen's daily lives with the increasing use of smartphones, which record the behaviors (e.g., dining, exercising, and shopping) of millions of people acting in different kinds of business venues. As a result, the LBSN data provide an alternative way to understand the crowd instead of delivering surveys to them. Moreover, compared with delivering surveys in several limited locations, the LBSN data make it possible to consider the business decision from a big picture of the whole city, rather than limited investigated areas.

In this paper, we mine the business opportunities from the LBSN data through recommending new business categories for a specific district, which guides the investors to decide to open what kind of business at which place. First, we partition a city into different districts, which helps to decide where to open a new shop. In particular, we define the properties of a business district and propose a greedy algorithm to discover business districts in a city. Furthermore, we define the business mining problem and propose the EmbeddingWARP algorithm to recommend new business categories for a specifically partitioned business district. We propose an embedding model [10] to learn category latent representations, which capture the functional correlations among different categories. Next, we model the interacting relations between a district and a category through the collaborative filtering technique. Then, we propose a Weighted Approximate-Rank Pairwise (WARP) loss [13] based ranking model to recommend business categories for a specific business district on the grounds of the learned category representations. To verify our system, we conduct experiments on Yelp data.

*The first author is an intern in Huawei Noah's Ark Lab when the work is done.

SIGIR '17, , August 07-11, 2017, Shinjuku, Tokyo, Japan

ACM ISBN 978-1-4503-5022-8/17/08…$15.00
http://dx.doi.org/10.1145/3077136.3080712

[1] https://www.franchisehelp.com/industry-reports/fast-food-industry-report/
[2] http://www.dollartimes.com/inflation/inflation.php

Experimental results show that our proposed method resolves the problem well.

We summarize the contributions as follows. 1) To the best of our knowledge, this is the first work mining the business opportunities for business owners based on LBSN data. Specifically, we define the business mining problem and propose a system to help business owners to discover commercial opportunities of new business categories from LBSNs. 2) We first define the business district property and propose a greedy algorithm to discover the business districts. This algorithm partitions a city into different districts, which helps to specify locations for business. 3) We propose a system for the business mining problem, which recommends new business categories for a specific business district based on the learned category embedding representations. Experimental results show that our proposed method successfully addresses the business mining problem.

2 RELATED WORK

LBSN applications. LBSNs play an important role in each citizen's daily lives with the increase of smartphones. Popular LBSNs such as Yelp and Foursquare, have attracted millions of users and hundreds of thousands of business shops. To improve user experience and prosper the businesses in LBSNs, a variety of new applications come out, e.g., point-of-interest (POI) recommendation and retail allocation system. On the one hand, through mining the user check-in data, POI recommendation systems [14, 15] are proposed to improve user experience, helping users to explore new places through the LBSNs. In addition, researchers mine the user check-in records and social preference to help make new friends in LBSNs [9, 12]. On the other hand, some researchers exploit the business data to serve the business owners in LBSNs, proposing the zone recommendation system [8], business prediction system [1], and retail allocation system [4, 6].

Connection with prior work. Our work employs the business data to mine the commercial opportunities of new shops for business owners in LBSNs. For mining the business opportunities, studies in [1, 4, 6, 8] involve two limitations. First, the studies in [1, 4, 6] consider a limited number of given locations. Although Lin et al. [8] consider a city as a whole, they recommend locations coarsely based on zones partitioned in the government city plan. Second, the work [4, 6] is designed for allocating chain stores, constrained in a specific categorical business type. The studies [1, 8] can cover more categories. Nevertheless, they recommend the business category based on the similarity, which ignores the correlations among business categories (e.g., partner-marketing relations) that have been proven to be important [2]. Therefore, methods in [1, 4, 6, 8] cannot resolve the business opportunity mining problem well.

3 PROBLEM DEFINITION

In the urban development, people usually gather for a specific target such as study, work, and travel. Different kinds of business venues are ensuingly open to satisfy the commercial needs of these people, which naturally forms a business district for a specific function, for instance, business districts around a university. From this phenomenon, we observe, 1) district forms based on a center location, e.g., the university, which contains a specific function attracting people gathering; 2) venues in a district are geographically adjacent to the center location. In the process of a business district's formation and expansion, new business shops are needed, which implies the business opportunities we are pursuing. Therefore, the key points of mining business opportunities constitute two steps: 1) discovering different districts, which guides where to open the business; 2) recommending the appropriate categories for a specific district, which guides to open what kind of business.

To sum up, we aim to predict what kind of businesses are needed for a partitioned business district. Before the formal problem definition, we define two basic terms as follows.

Definition 1 (Business Venue). *A business venue is a shop v, with geographical latitude and longitude $\langle lat, lon\rangle$, and a category set C_v containing categories labeled to the venue v.*

Definition 2 (Business District). *A business district V_d is an aggregation of different business venues, in which venues are geographically adjacent to each other.*

The business mining problem aims to discover emerging business venue categories in a business district V_d. In practice, a business venue v contains at least one category type. Thus, a category set C_v contains at least one element. In addition, a business district V_d contains several venues, which may consist of the same or different categories. Therefore, a business district V_d contains a series of categories, forming a set C_V. Each element $c \in C_V$ appears in the district V_d with different times counting based on the number of venues. Moreover, we assume that when a category attracts more shops, this category indicates bigger commercial values. Now, the business mining problem can be formally defined as follows.

Definition 3 (Business Mining Problem). *The business mining problem aims to recommend a category set C_R for a business district V_d, given existing category set C_V and corresponding appearing times of each category.*

4 METHOD

4.1 Discovering Business Districts

We propose a greedy business district discovering algorithm. The discovered district should satisfy the following two requirements: 1) the venues in the district are geographically adjacent to the center venue; 2) the district should contain the business center in the local, which is the center location playing the role of gathering crowds. To discover the business districts, we first define the geographical location graph, which captures the venues' geographical topological properties.

Definition 4 (Geographical Venue Graph). *Given venues in a city as a vertex set V, the geographical venue graph G is constructed by generating edges among any two vertex pair $\langle v_i, v_j\rangle$, whose geographical distance is less than a distance threshold.*

Based on the defined venue graph G, the business districts discovering algorithm aims to find a graph cutting method satisfying the two requirements. To satisfy requirement 2, we need to find the business center first. As the LBSNs attract millions of people to record their daily lives via check-in activities, the number of check-ins reflects a business venue's ability to attracting the crowd.

Hence, we define the landmark venue as the business center in the local as follows.

Definition 5 (Landmark Venue). *A landmark venue v_l is the most checked-in location in a business district V_d.*

With Definition 4 and Definition 5, the business district discovering algorithm could be elucidated as follows. Denote $\mathcal{V}_c$ as the set of discovered districts, then the target of our business district discovering algorithm is to find $\mathcal{V}_c$, given the geographical venue graph G. First, we sort all the venues according to the check-in frequency to ensure landmark venue locating in the local business district in later operations. Next, we visit venues with the sorted order and treat the visited venue as landmark venue to construct a business district. We employ the breadth-first search (BFS) to visit the geographical venue graph G and select the venues whose distances to the landmark venue are less than the threshold d_t. As a result, the landmark venue and the selected venues constitute a business district V_d.

4.2 EmbeddingWARP Model

After partitioning a city into different districts, mining business opportunities aims to predict emerging business categories for a specific district by the correlations among different categories. To address this task, we propose the EmbeddingWARP algorithm, in which we first learn the category representations and then recommend new categories for a specific district through a WARP loss based ranking model.

4.2.1 Learning Category Representations. To capture the correlations of different categories, we propose an embedding model to learn the category representations. Our model is based on two assumptions: 1) categories belonging to the same venue are highly correlated, and 2) adjacent venues in the geographical venue graph (Definition 4) are highly correlated. Borrowing the idea from network embedding [3, 10], we propose the category embedding learning model. To satisfy assumption 1, we use the categories labeled the same venue to represent the venue together through an average operator; furthermore, to satisfy assumption 2, we use the neighbors in the geographical venue graph to represent each venue.

Formally, we formulate the category embedding model as follows. Supposing that $V_n = \{v_1, \ldots, v_k\}$ denotes the neighbors of a venue v_i in the venue graph G. Then, the objective function is to maximize the following log probability,

$$\sum_{v_i \in G} \sum_{v_j \in V_n} \log \Pr(v_j | v_i). \tag{1}$$

Here, the probability $\Pr(v_j|v_i)$ is formulated using a softmax function. Denote $\mathbf{v}_j$ and $\mathbf{v}_i$ are the latent vector representations of adjacent venue v_j and target venue v_i, respectively. Then the probability function $\Pr(v_j|v_i)$ can be defined as,

$$\Pr(v_j|v_i) = \frac{\exp(\mathbf{v}_j \cdot \mathbf{v}_i)}{\sum_{v_i \in V} \exp(\mathbf{v}_j \cdot \mathbf{v}_i)}, \tag{2}$$

where V is the venue set and $(\cdot)$ is the inner product operator.

In order to efficiently learn the objective function, we adopt negative sampling to improve the optimization [10]. For each venue v_i, we sample a set of K venues V_s that are not adjacent to v_i. Moreover, we leverage the category embedding representations to represent the venue vector. We assume each venue v_i is labeled a set of categories C_i. Suppose $\mathbf{c}_i$ denotes the embedding of a category c_i, we denote each venue vector $\mathbf{v}_i = \frac{1}{|C_i|} \sum_{c_i \in C_i} \mathbf{c}_i$. Therefore, the objective function of the category embedding model is as follows,

$$\sum_{v_i \in G} \sum_{v_j \in V_n} (\log \sigma(\mathbf{v}_j \cdot \mathbf{v}_i) + \sum_{v_s \in V_s} \log \sigma(\mathbf{v}_s \cdot \mathbf{v}_i)). \tag{3}$$

4.2.2 Business Recommendation Model. Treating each business district as a "user", the business category as an "item", and the appearing times of each category as the "rating", the business mining problem can be formulated as a traditional recommendation problem. We are encouraged to employ collaborative filtering methods to recommend a category to the target district if the category appears in some "similar" districts but not in the target district. Thus, we propose a WARP loss [13] based latent ranking model to recommend top-N new categories for a specific district.

We use $a_{c,d}$ to denote the occurrence number of a category c at a district V_d. If the category does not appear in the district, we set $a_{c,d} = 0$. We assume that the bigger the occurrence number $a_{c,d}$ is, the category c is more appropriate for the district V_d. Moreover, we use $\mathbf{d}_d$ to denote the latent representation for district V_d. The appearance of a category c at a district V_d is modeled by $\phi_{c,d} = \mathbf{d}_d \cdot \mathbf{c}_c$, namely dot product of district and category latent representations. Thus, for a category c appearing in a district d with $a_{c,d}$, the pairwise order can be measured as follows,

$$rank(c, d) = \sum_{c' \in C} I(a_{c,d} > a_{c',d}) I(\phi_{c,d} < \phi_{c',d} + \epsilon), \tag{4}$$

where $I(\cdot)$ is an indicator function and ϵ is a positive margin number. For each district V_d, we assume the appearing categories in this district as C_d^+. Thus, the WARP loss is defined as follows,

$$L_{warp} = \sum_{d \in \mathcal{V}_D} \sum_{c \in C_d^+} L(rank(c, d)), \tag{5}$$

where $L(r) = \sum_{i=1}^{r} \frac{1}{i}$ is a function converting the ranking order into a loss. In this objective function, the category representation is derived in the category embedding model, so our target is to learn the district representations. To avoid overfitting we constrain the district latent vector in a ball, $||\mathbf{d}_d||_2 \leq S$, for all $V_d \in \mathcal{V}_D$.

We use the learning scheme in [7] to learn the objective function. Specifically, we iterate through each category c appearing in a district d and sample a negative category c' to update the vectors. Here, we use Bootstrap technique to generate samples. We assume we get the negative category c' until the M^{th} sample, satisfying $(a_{c,d} - a_{c',d})(\phi_{c',d} + \epsilon - \phi_{c,d}) > 0$. Then, we use the stochastic gradient descent (SGD) algorithm to learn the latent factors of districts. Denote $\delta(x) = \sigma(x) * (1 - \sigma(x))$, where $\sigma(\cdot)$ is the sigmoid function. The gradient of L_{warp} with respect to the latent factor of district V_d is as follows:

$$\frac{\partial L_{warp}}{\partial \mathbf{d}_d} = L(\lfloor \frac{|C|}{M} \rfloor) \delta(\phi_{c',d} + \epsilon - \phi_{c,d})(\mathbf{c}_c - \mathbf{c}_{c'}). \tag{6}$$

The latent factor is updated as : $\mathbf{d}_d = \mathbf{d}_d + \eta \frac{\partial L_{warp}}{\partial \mathbf{d}_d}$, where η is the learning rate. After the update, we project the weights to satisfying the constraint $||\mathbf{d}_d||_2 \leq S$, $\mathbf{d}_d = \mathbf{d}_d \frac{||\mathbf{d}_d||_2}{S}$.

Once latent representations of districts have been learned, we measure the appropriateness of a category c at a district V_d by

$\phi_{c,d} = \mathbf{d}_d \cdot \mathbf{c}_c$. We recommend the business categories by sorting the candidate categories in descending order of the predicted scores and selecting N top-ranked categories.

5 EXPERIMENTS

In this section, we conduct experiments on the LBSN data from the Yelp dataset challenge 2015. We extract information about local businesses in four cities, including Pittsburgh, Charlotte, Phoenix and Las Vegas. To recommend meaningful categories, we ignore the parent category label in a group of category labels for a venue. For example, when a business venue is labeled with "Sandwiches and Restaurants", we ignore the "Restaurants" and adopt "Sandwiches" as the target category for the venue.

5.1 Experimental Settings

We handle the business mining problem with two steps: discovering business districts and recommending categories for a specific district. For discovering business districts, we set the distance threshold as one kilometer. For the discovered districts, we filter the districts with less than five venues, because those sporadic business venues do not compose real districts. Then, for the selected districts, we randomly choose 80% venues for training and the remaining 20% for testing. We use general metrics in recommendation systems, **precision** and **recall**, to evaluate our system.

After discovering business districts, recommending new categories for a district is a typical recommendation problem. Treating the occurrence number of a category in a district as one kind of implicit feedback, we can recommend categories for a district using state-of-the-art recommendation methods. Therefore, to demonstrate the effectiveness of our proposed method, we compare the proposed model with the following baseline methods, **Random Selection**, **Most Popular**, **BPRMF** [11], and **WRMF** [5].

5.2 Experimental Results

We partition the city into different districts and generate the discovered districts for each city shown in Table 1. Compared with the official neighborhoods partition, our partition is more fine-grained. For instance, the city of Phoenix contains 15 urban villages in official[3], while, Phoenix contains 364 districts according to our partitioning algorithm.

Table 1: Statistics of discovered districts

City	Las Vegas	Phoenix	Charlotte	Pittsburgh
#Districts	367	364	222	85

After discovering business districts, recommending categories for a district is a classic recommendation problem. Compared with general recommendation methods, the proposed EmbeddingWARP model captures the correlations among categories, which makes our proposed algorithm achieve the best performance.

City	Model	P@1	P@5	P@10	R@1	R@5	R@10
Las Vegas	Random	0.03	0.02	0.02	0.003	0.009	0.02
	MostPop	0.16	0.15	0.13	0.03	0.10	0.16
	BPRMF	0.22	0.19	0.17	0.023	0.96	0.17
	WRMF	0.33	0.25	0.20	0.054	0.15	0.22
	EmbeddingWARP	**0.40**	**0.33**	**0.30**	**0.07**	**0.20**	**0.33**
Phoenix	Random	0.02	0.02	0.02	0.01	0.02	0.03
	MostPop	0.15	0.13	0.11	0.04	0.13	0.20
	BPRMF	0.16	0.16	0.14	0.03	0.13	0.24
	WRMF	0.22	0.16	0.13	0.05	0.12	0.22
	EmbeddingWARP	**0.32**	**0.23**	**0.24**	**0.09**	**0.18**	**0.30**
Charlotte	Random	0.01	0.02	0.02	0.002	0.009	0.03
	MostPop	0.10	0.11	0.10	0.03	0.14	0.20
	BPRMF	0.23	0.15	0.12	0.04	0.13	0.20
	WRMF	0.23	0.17	0.14	0.04	0.16	0.26
	EmbeddingWARP	**0.30**	**0.24**	**0.20**	**0.08**	**0.21**	**0.33**
Pittsburgh	Random	0.03	0.02	0.02	0.007	0.03	0.05
	MostPop	0.16	0.16	0.12	0.04	0.20	0.30
	BPRMF	0.21	0.17	0.16	0.05	0.20	0.31
	WRMF	0.24	0.15	0.14	0.04	0.11	0.20
	EmbeddingWARP	**0.29**	**0.20**	**0.19**	**0.07**	**0.21**	**0.33**

Figure 1: Business recommendation results

6 CONCLUSION AND FUTURE WORK

In this paper, we propose the business mining problem on LBSN data, which discoveries commercial opportunities of new shops for business owners. First, we define the properties of a business district and propose a greedy algorithm to partition a city into different districts. Then, we propose an embedding learning based rank model to predict emerging categories for a specific district. In the future, we will employ knowledge from other cities to boost the recommendation performance using transfer learning.

7 ACKNOWLEDGEMENT

This work was partially supported by the Research Grants No. CUHK 14203314 and No. CUHK 14205214 of the General Research Fund, and 2015 Microsoft Research Asia Collaborative Research Program (Project No. FY16-RES-THEME-005).

REFERENCES

[1] O. Al Sonosy, S. Rady, N. L. Badr, and M. Hashem. 2015. Exploiting location based social networks in business predictions. In *International Conference on Innovations in Information Technology*.

[2] J. Bao, A. Deshpande, S. Mcfaddin, and C. Narayanaswami. 2014. Partner-marketing using geo-social media data for smarter commerce. In *Ibm Journal of Research and Development*, Vol. 58. 6:1–6:12.

[3] Steven Skiena Bryan Perozzi, Rami Alrfou. 2014. DeepWalk (online learning of social representations). In *KDD*.

[4] Bo Hu. 2014. *Recommendation in Location-based Social Networks.* Ph.D. Dissertation. Applied Sciences:.

[5] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *ICDM*.

[6] Dmytro Karamshuk, Anastasios Noulas, Salvatore Scellato, Vincenzo Nicosia, and Cecilia Mascolo. 2013. Geo-spotting: Mining Online Location-based Services for Optimal Retail Store Placement. In *KDD*.

[7] Xutao Li, Gao Cong, Xiao-Li Li, Tuan-Anh Nguyen Pham, and Shonali Krishnaswamy. 2015. Rank-GeoFM: A Ranking Based Geographical Factorization Method for Point of Interest Recommendation. In *SIGIR*.

[8] Jovian Lin, Richard J Oentaryo, Ee-Peng Lim, Casey Vu, Adrian Vu, Agus T Kwee, and Philips K Prasetyo. 2016. A Business Zone Recommender System Based on Facebook and Urban Planning Data. In *Advances in Information Retrieval*. Springer.

[9] Haiping Ma, Huanhuan Cao, Qiang Yang, Enhong Chen, and Jilei Tian. 2012. A habit mining approach for discovering similar mobile users. In *WWW*. ACM.

[10] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *NIPS*.

[11] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*. AUAI Press.

[12] Salvatore Scellato, Anastasios Noulas, and Cecilia Mascolo. 2011. Exploiting place features in link prediction on location-based social networks. In *KDD*.

[13] Jason Weston, Samy Bengio, and Nicolas Usunier. 2010. Large scale image annotation: learning to rank with joint word-image embeddings. *Machine learning* (2010).

[14] Shenglin Zhao, Tong Zhao, Irwin King, and Michael R Lyu. 2017. Geo-Teaser: Geo-Temporal Sequential Embedding Rank for Point-of-interest Recommendation. In *WWW*. 153–162.

[15] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R Lyu, and Irwin King. 2016. Stellar: spatial-temporal latent ranking for successive point-of-interest recommendation. In *AAAI*.

[3] https://en.wikipedia.org/wiki/Phoenix, _Arizona