2581-4575

International Journal For Recent Developments in Science & Technology

Crossref

A Peer Reviewed Research Journal

COMPARISON OF MULTIPLE APPS AND POPULARITY PREDICTION BASED ON USER REVIEWS

[1] M.SESHUBHAVANI, [2] K.SRI KRISHNA DATTU, [3] J.SUJEEVA RAO, [4] CH.YASHWANTH

[1]ASSOCIATE PROFESSOR, CSE DEPARTMENT, NRI INSTITUTE OF TECHNOLOGY, POTHAVARAPPADU

[2, 3, 4] UG STUDENT, CSE DEPARTMENT, NRI INSTITUTE OF TECHNOLOGY, POTHAVARAPPADU

Abstract: In recent years, bike-sharing systems are wide deployed in many Brobdingnagian cities that provide a cost-efficient fashion. With the prevalence of bike-sharing systems, plenty of companies are a vicinity of the bike sharing market, leading to additional and additional fierce competition. To be competitive, bike-sharing companies and app developers have to be compelled to build strategic picks for mobile apps development. Therefore, it is vital to predict and compare the recognition of varied bike-sharing apps. However, existing works for the most part focus on predicting the recognition of one app, the popularity contest among altogether completely different apps has not been explored but. In this paper, we've a bent to aim to forecast the recognition contest between Mobike and Ofo, a pair of most well-liked bike-sharing apps in China. We've a bent to develop CompetitiveBike, a system to predict the recognition contest among bike-sharing apps. Moreover, we've a bent to conduct experiments on real-world datasets collected from eleven app stores and Sina Weibo, and so the experiments demonstrate the effectiveness of our approach.

Keywords: Bike-sharing app, Mobile app, Competitive analysis, Popularity prediction

1 Introduction:

In recent years, shared transportation has big enormously, that provides US an economical style. Among the assorted sorts of shared transportation, public bike-sharing systems [1], [2], [3] are wide deployed in several metropolitan areas (e.g. NY town within the US and national capital in China). A bike-sharing system provides short-run bike rental service with several bicycle stations distributed in a city [4]. A user will rent a motorbike at a close-by bike station, and come back it at another bike station close to his/her destination. The worldwide prevalence of bike-sharing systems has galvanized millions of active analysis, like bike demand prediction [5], [6], [7], bike rebalancing improvement [8], and bike lanes designing [9].

More recently, station-less bicycle-sharing systems are getting the thought in several huge cities in China like national capital and Shanghai. Mobike1 and Ofo2 square measure 2 hottest station-less bicycle-sharing systems. In contrast to ancient public bike-sharing systems, station-less bike sharing systems aim to resolve "the last one mile" issue for users. Exploitation the Mobike/Ofo mobile app, users can search and unlock near bikes. Once users attain their destinations, they do not need to come back the bikes to the selected bike station. Instead, they can park the bicycles at a location additional convenient for them. Therefore, it's easier for users to rent and come back bikes than ancient bike-sharing systems.

As bike-sharing apps become more and more common, loads of firms are part of the bike-sharing market, resulting in fierce competition. To thrive during this competitive market, it's very important for bike-sharing firms and app developers to know their competitors and so create strategic choices consequently [10] for mobile app development and evolution [11]. Therefore, it's important and necessary to predict and compare the long run quality of various bike-sharing apps.

When users transfer and install a mobile app, they'll submit user expertise to the app store [12], [13], [14]. Specifically, users could transfer their requirements (e.g. useful requirements), preferences (e.g. UI preferences) or sentiment (e.g. positive, negative) through reviews, further as their satisfaction level through ratings. On-line social media is differently to share the user expertise of a mobile app. once users really use the bike, they'll share the ride expertise on social media. Specifically, users could record the sensation of the ride, the benefits and drawbacks of the bike/system, or the comparison with different bikes/systems. Each users' on-line and offline expertise can have an effect on the popularity of the apps, thereby moving their contest outcome. Therefore, app store knowledge and microblogging knowledge square measure complementary, and can describe a mobile app from totally different views. During this paper, we have a tendency to study the problem of competitive analysis and recognition prediction of bike-sharing apps using app store knowledge and microblogging knowledge.

To the most effective of our information, the matter of predicting the fight of mobile apps has not been well investigated within the literature. There square measure many challenging inquiries to be answered. The way to estimate and forecast the recognition contest outcomes of bike-sharing apps? The way to extract effective options to characterize the recognition of bike-sharing apps from multi-source data?

To answer these queries, we have a tendency to propose CompetitiveBike, a system that predicts the outcomes of the recognition contest among bike-sharing apps investment app store knowledge and microblogging knowledge. We have a tendency to 1st get app descriptive statistics and sentiment data from app store knowledge, and descriptive statistics and comparative data from microblogging knowledge. Exploitation these knowledge, we extract both coarse-grained and fine-grained competitive options. Finally, we train a regression model to predict the outcomes of recognition contest. We have a tendency to create the following contributions.

International Journal For Recent Developments in Science & Technology

2581-4575

Crossref

A Peer Reviewed Research Journal

(1) This work is that the 1st to review the matter of competitive analysis of bike-sharing apps. We have a tendency to use 2 indicators for the comparison: i) competitive relationship to point that app is additional popular; and ii) competitive intensity to measure the recognition gap between the 2 apps/systems.

(2) To predict contest between apps, we have a tendency to extract options from different views as well as the descriptive data of apps, users' sentiment, and comparative opinions. Exploitation the fundamental data, we have a tendency to additional extract 2 novel options: coarse-grained and fine-grained competitive features, and opt for Random Forest for prediction.

(3) To Gauge CompetitiveBike, we have a tendency to collect knowledge regarding Mobike and Ofo from 11 app stores and Sina Weibo. With the info collected, we have a tendency to conduct intensive experiments from totally different views. we discover that the Random Forest model performs well on competitive relationship prediction (the Accuracy is seventy one.4%) as well as competitive intensity prediction mix of the coarse-grained and fine-grained competitive options improves performance in contest prediction, and a mix of knowledge from app store and microblogging conjointly improves performance in contest prediction. The results demonstrate the effectiveness of our approach.

2 Related Work

2.1 Popularity Prediction:

Recently, a major effort has been spent on predicting quality of mobile Markov Model (PHMM) to model the recognition data of mobile apps. [16] Projected a stratified model to forecast the app downloads. Malmi [17] found that there existed association between app quality and therefore the found that there's a powerful correlation between rating and therefore the downloads.

Our work differs from and probably outperforms the previous add many aspects. First, we have a tendency to target the matter of competitive analysis and recognition prediction of bike-sharing apps, rather than the prediction of one app. Second, we have a tendency to predict the recognition contest investment multi-source information (i.e., app store information and microblogging data) that area unit typically complementary and might mirror.

2.2 Competitive Analysis:

Competitive associate analysis involves the first identification of potential risks and opportunities to assist managers creating strategic choices for an enterprise [10]. [19] elect subjective sentences from reviews which debate common the social media of the 3 largest dish chains, and therefore the results unconcealed the generative model for comparative sentences, collectively modeling 2 levels of comparative relations: the amount of sentences and therefore the level of entity pairs. al. [22] projected to scan reviews to update a product comparison network. These studies conduct competitive

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

analysis merely via linguistics analysis of in distinction, our work extracts options from totally different views including the descriptive info of apps, user's sentiment, and comparative opinions. Exploitation the fundamental info, we tend to more extract coarse-grained and finegrained competitive options, and train a model to predict contest.

3 Problem Statement and System framework

3.1 Problem Statement:

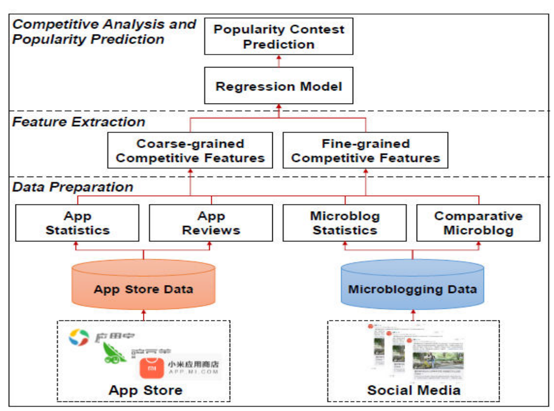
The problem are often expressed as follows: given the app store knowledge and microblogging data concerning Mobike and Ofo, we wish to predict that app are additional widespread galvanized by [23], the recognition of Mobike (or Ofo) are often measured by the downloads, and also the contest (P C) between Mobike and Ofo are often outlined by the distinction in their downloads Dm between Mobike and Ofo are often one amongst the 2 possibilities: 1) Mobike is additional popular than Ofo, or 2) Ofo is additional widespread than Mobike. In line with Formula (1), once P C > zero, Mobike is additional popular; otherwise, Ofo is additional widespread. Mobike and Ofo is that the definite quantity of P C. Formally, we tend to extract feature set X from app store knowledge and microblogging data, then we wish to predict the recognition contest Y . Y = , given X(1:t 1)(=) and Y our objective is to predict Y (t 1) The summary of the framework is illustrated in Figure two, that primarily consists of 3 layers: knowledge preparation, feature extraction, and competitive analysis we tend to get app statistics and reviewers' sentiment from app store knowledge, and microblogging statistics and comparative info from To effectively extract and quantify the factors impacting mobile app contest, we tend to extract options from totally different views as well as the inherent descriptive info of apps.

3.2 System Framework

The summary of the framework is illustrated in Figure a pair of that chiefly consists of 3 layers: information preparation, feature extraction, and competitive analysis and popularity prediction.

Data Preparation: we tend to acquire app statistics and reviewers' sentiment from app store information, and microblogging statistics and comparative info from microblogging information.

Feature Extraction: To effectively extract and quantify the factors impacting mobile app contest, we tend to extract options from totally different views together with the inherent descriptive info of apps, users' sentiment, and comparative opinions. With this in



of, we tend to more extract 2 novel sets of options: coarse-grained and fine-grained competitive feat user.

Competitive Analysis and recognition Prediction:

With these 2 extracted feature sets, we tend to train a model to predict the recognition contest between Mobike and Ofo.

4 Popularity contest Prediction

In this section, we tend to 1st analyze the factors impacting the recognition contest between Mobike and Ofo, then extract coarse-grained and fine-grained competitive features from these factors to characterize contest. Finally, we train a model to predict contest.

4.1 Coarse grained Competitive Features:

Once users transfer and install a mobile app, they may submit reviews and ratings to the app store. As an example, a user wrote: "The Mobike app cannot launch these days; it had been still okay yesterday, what's the matter? It's terrible! " per the review, we tend to believe that app store information (e.g. reviews, ratings) will replicate users' on-line expertise with the app. typically, users might transfer their needs (e.g. purposeful requirements), preferences (e.g. UI preferences), or sentiment (e.g. positive, negative) through reviews, and they may conjointly rate the app supported their overall satisfaction. Therefore, we extract options from reviews and ratings to characterize contest.

App Statistics, Generally, the numerical statistics of reviews and ratings in each time window will replicate the recognition of the app. In different words, a bigger number of reviews and a better rating score might indicate that the app is a lot of popular. We tend to use the distinction between app's review variety DN (and rating scores DS) to characterize contest. A tiny low price of DN (and DS) indicates that they need similar variety of reviews (and rating score), thus their competition is a lot of intense. **Sentiment Similarity**. Besides numerical statistics, app reviews will categorical users' sentiment. We tend to use a Chinese sentiment analyzer known as SnowNLP5 to analyze the sentiment of reviews. We tend to calculate the sentiment price si of each review at time instant ti, then we tend to get the sentiment distribution vector VI = (p1, p2, p3) at time ti, where p1, p2, p3 is corresponds to negative, neutral and positive sentiment proportion severally. The extracted sentiment sequences square measure just for one app, after we take into account the competition between 2 apps, we tend to cipher sentiment similarity to capture the distinction of users' sentiment concerning these apps, and also the similarity will be measured by hard the circular function similarity [24].

The upper similarity means that those users' opinions concerning them square measure a lot of similar, and also the competition between them is a lot of intense options from Microblogging. Once users ride the bike of various apps, they may share their riding expertise on social media. Associate example of a micro blog is like this: "This is my 1st ride of Mobike, it's therefore cool!" we tend to believe online social media is in a different way to precise users' riding expertise. Therefore, we extract options from microblogging information to assist perceive the recognition contest of various apps. Microblogging Statistics. Within the "Mobike & Ofo" dataset, the quantity of micro blogs, users, reposts, comments, and likes will replicate the eye concerning Mobike and Ofo on microblogging, the larger price indicates a lot of intense competition between Mobike and Ofo.

In the "Mobike" dataset, a lot of micro blogs that contain the keyword "Ofo" imply that Ofo is a lot of often mentioned within the "Mobike" dataset. We use the ratio (Rom) of "Ofo" and "Mobike" to characterize the competition. Formally, Rom = MNo / MNm, wherever MNo and MNm represent the quantity of micro blog that contains "Ofo" and "Mobike", severally. Similarly, within the "Ofo" dataset, we use the quantitative relation (Rmo) of "Mobike" and "Ofo" to characterize the competition. The higher ratios, the lot of intense competition.

4.2 Fine-grained Competitive options

Each coarse-grained competitive feature may be a statistic with time window of one week. In whenever window, we have a tendency to extract the temporal dynamics of the coarse-grained competitive options because the fine-grained competitive options to characterize the trend of the sequence [25]. Overall Descriptive Statistics describe the essential properties of the coarse grained competitive options from multiple aspects. we have a tendency to extract the mean, variance, median, minimum and most as options.

Hopping Counts will effectively describe the "pulse" of sequence and is calculated because the range of parts whose values square measure bigger than their next element. This feature is employed to characterize the fluctuation of the sequences. Lengths of Longest Monotonous Subsequences describe the dimensions of gradient descent or ascent patterns during a sequence. We have a tendency to examine the longest monotone (including increasing and decreasing) subsequences, and use the lengths of those two subsequences to explain the tendency of the sequence.

4.3 Popularity Contest Prediction:

With these 2 extracted feature sets, we wish to predict the recognition contest in the future; we have a tendency to use regression-based ways. Since the extracted options square measure sequences, and therefore the time window is one week, we have a tendency to treat sequential many weeks as the coaching set, and then compare the progressive regression models. Section 6 has the small print on the models we have a tendency to compare and therefore the one we have a tendency to eventually use.

5 Conclusion

In this paper, we have a tendency to specialize in the matter of competitive analysis and recognition prediction over Mobike and Ofo. We have a tendency to propose CompetitiveBike to predict the popularity contest between Mobike and Ofo leverage app store information and microblogging information. Specifically, we have a tendency to 1st extract options from completely different views together with the inherent descriptive info of apps, users' sentiment, and comparative opinions. With this info, we have a tendency to any extract 2 sets of novel options: coarse-grained and fine-grained competitive features. Finally, we choose the Random Forest formula to predict the recognition contest. Moreover, we have a tendency to collect information regarding 2 bike-sharing apps from eleven on-line mobile app stores and Sina Weibo, implement in depth experimental studies, and also the results demonstrate the effectiveness of our approach. In the future work, we are going to enrich our drawback statement and system framework by learning from the classical economic theories on competitive analysis [26], [27]. So as to supply competitive analysis for mobile apps, we are going to read the mobile apps competition as a long event, and generate the event plot [28] and gift descriptive info concerning contest to enrich the competitive analysis. Besides, we are going to improve the prediction model by analyzing the couplings [29], [30] among options and deciding their mutual influence. Moreover, we are going to collect a lot of classes of apps to counterpoint our datasets, and extend the generality of our approach to alternative apps.

6 References

1. B. Guo, Y. Ouyang, C. Zhang, J. Zhang, Z. Yu, D. Wu, and Y. Wang, "Crowdstory: Fine-grained event storyline generation by fusion of multi-modal crowdsourced data," Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 1, no. 3, p. 55, 2017.

2. L. Cao, Y. Ou, and S. Y. Philip, "Coupled behavior analysis with applications," IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 8, pp. 1378– 1392, 2012.

3. L. Cao, T. Joachims, C. Wang, E. Gaussier, J. Li, Y. Ou, D. Luo, R. Zafarani, H. Liu, G. Xu et al., "Behavior informatics: A new perspective," IEEE Intelligent Systems, vol. 29, no. 4, pp. 62–80, 2014.

4. Z. Zhang, C. Guo, and P. Goes, "Product comparison networks for competitive analysis of online word-of-mouth," ACM Transactions on Management Information Systems (TMIS), vol. 3, no. 4, p. 20, 2013.

5. E. Malmi, "Quality matters: Usage-based app popularity prediction," in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. ACM, 2014, pp. 391–396.

6. P. DeMaio, "Bike-sharing: History, impacts, models of provision, and future," Journal of Public Transportation, vol. 12, no. 4, p. 3, 2009.

7. S. Shaheen, S. Guzman, and H. Zhang, "Bikesharing in Europe, the Americas, and Asia: past, present, and future," Transportation Research Record: Journal of

2581-4575

International Journal For Recent Developments in Science & Technology

Crossref

A Peer Reviewed Research Journal

the Transportation Research Board, no. 2143, pp. 159–167, 2010.

8. J. Pucher, J. Dill, and S. Handy, “Infrastructure, programs, and policies to increase bicycling: an international review,” Preventive medicine, vol. 50, pp. S106–S125, 2010.