Computers and Electrical Engineering 000 (2017) 1–14

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

# Complex network measurement and optimization of Chinese domestic movies with internet of things technology☆

Jia Xiao[a,*], Xin Li[a], Shanzhi Chen[b], Yan Wang[c], Jiaxing Han[a], Zhili Zhou[d]

[a] State Key Laboratory of Networking and Switching, Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Haidian District, Beijing 100876, China
[b] State Key Laboratory of Wireless Mobile Communications, China Academy of Telecommunications Technology, Beijing 100083, China
[c] Xinyang Power Supply Company, Henan Electronic Power Company, Xinyang 464000, Henan, China
[d] Jiangsu Engineering Centre of Network Monitoring & School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

ARTICLE INFO

*Article history:*
Received 3 June 2017
Revised 8 July 2017
Accepted 11 July 2017
Available online xxx

*Keywords:*
Semantic feature
Movie complex network
Community detection
Genetic algorithm

ABSTRACT

The market performance prediction of domestic motion picture is an important problem that is worthy of study. In this paper, by incorporating Chinese fine-grained semantic features, we propose a method of community detection and genetic optimization especially for Chinese domestic films. These semantic features, also named as gene elements, are used as nodes to construct a movie complex network. Through leveraging the influence of the node both in the whole network and in the internal community, four unique communities are revealed for successful Chinese movies. Then the Genetic Algorithm (GA) with a proposed novel fitness function is used to obtain the optimal cluster of gene elements. For the other operations in GA (i.e. initialization, selection, crossover and mutation), the parameters are also be optimized by a distinctive evaluation method. Finally, the experiments on the data of Chinese motion pictures in 2016 demonstrate the efficacy and accuracy of the overall system.



## 1. Introduction

In recent years, the Chinese multimedia industry, especially the film industry, has great development. From New Year's Day of 2016 to February 23rd, the total box-office revenues surpassed 10 billion. For example, the box-office revenue of a domestic movie "Mermaid", directed by Stephen Chow, had been smashing historic record since it was shown. This most dynamic market is attracting more and more investment capital. However, due to the box office of different film differs greatly, to get a stable investment profit is difficult.

Fig. 1 displays the film box-office in low-to-high sorted order. The horizontal axis X depicts the number of films on screen from 2015 to May 2016, while the vertical coordinate Y represents the box-office receipt of each film. Based on the revenues data retrieved from the website named 'M1905' with help of jsoup crawler, we found that almost one fourth of films occupy a large cut of the market, which means that the probability of a box-office hit is only about 25% vice versa.

With this observation, there are two problems which are worthy of study. Firstly, which features domestic films with high box-office shall have. Secondly, whether it is possible to forecast a set of features by which a successful Chinese domestic

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. Zhihan Lu.

* Corresponding author.
*E-mail address:* spaceman18@126.com (J. Xiao).

http://dx.doi.org/10.1016/j.compeleceng.2017.07.009
0045-7906/

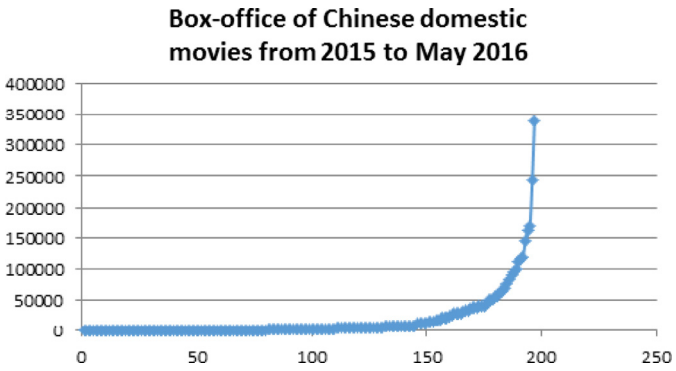


**Fig. 1.** Receipts distribution curve.

movie can be produced. In this paper, one hundred million yuan revenue is set as the criterion of high receipt from a commercial standpoint.

We discuss a challenging problem of predicting the potential "successful" (high box-office receipts) characteristics of Chinese domestic motion pictures. Most research works mainly focus on both traditional factors (cast, producer, issuer, genre, award, etc.) and online signals (web search counts, comment counts, view times, like numbers, dislike numbers, etc.), few of them have taken the semantic resource of films into consideration. Therefore, using semantic resource is a novel approach to discover potential hit films in Chinese market.

The rest of this paper is organized as follows. Section 2 presents related work on analyzing movie itself and market performance. In Section 3, we briefly introduce the core concepts utilized in this paper. In Section 4, we describe the construction of Chinese domestic movie complex network. Sections 5 and 6 present the complex network measurement of domestic movie features and genetic optimization. Finally, Section 7 summarizes this paper.

## 2. Related work

There already exist a great number of works with various methods on box office prediction. Most of works built analytical models based on two distinct types of factors. One is the set of classical attributes. The other is users' anticipations and responses collected from various online platforms.

References [1–10] proposed to predict market responses with the help of internet user behavior, such as web search clicks, reviews collected from social networks and Wikipedia, and indexes on video websites etc. These user behaviors could be processed together in diversified ways, some of which are popular approaches, e.g. correlation analysis, sentiment mining and machine learning. Furthermore, some papers even try to combine these online resources with the offline data from theatres, to solve the problems of data deficiency, distortion, or pollution.

Some researchers studied the relationship between market performance and traditional elements in references [11–15], including actor, director, budget, script, award, genre, even sequel and so forth, where they mainly resort to correlation analysis. However, they have not drawn a common conclusion. From our perspectives, we think that multiple factors play an integrated role as so-called content while each of them weights distinctively to different films. Jinni had gone much further than all the other video competitors by taking a taxonomic approach, cataloging titles, analyzing user reviews and metadata as presented in reference [16]. This novel technology is also used to power semantic discovery engines for movies and TV shows.

It should be noticed that Chinese movie market is also studied by using machine learning and correlation analysis on online and offline data. Reference [8] proposed a method of utilizing Tencent microblog to predict the box-office revenues of movies screened in China. References [12–13] analyzed the relationship between the receipts and classical elements of movies released in China. Reference [17] applied the algorithms of pruned random forest to analyze data from theatres to predict box office receipts. Reference [18] discussed the relationship between WOM (online word of mouth) and the box-office of foreign films in China.

From all the papers listed above, we could get three conclusions. First, nearly none of them specifically discussed Chinese domestic films market which is featured by rapid growth in these years as shown in Fig. 2.

Secondly, the issues that the papers studied Chinese movie market concerned are still subject to further research. For example, Tencent microblog had already withdrawn from the market, so the method introduced in the reference [8] is not applicable. WOM could not be retrieved before releasing, thus the method provided in the reference [18] can not be used for revenues prediction.

Last but not the least, box-office is determined by multiple factors which are interrelated and interact on each other. Therefore, the method of linear regression based on single factor, adopted by references [12–13], is inappropriate. Moreover, as mentioned in references [19], lacking of unified classification standards for all films could also lead to inaccurate prediction.

Please cite this article as: J. Xiao et al., Complex network measurement and optimization of Chinese domestic movies with internet of things technology, Computers and Electrical Engineering (2017), http://dx.doi.org/10.1016/j.compeleceng.2017.07.009