

## Evaluation And Customization of TV Drama

Yi Li

North China Electric Power University, Baoding 071003, China

18331129603@163.com

**Keywords:** TV drama; Grey system theory; Model

**Abstract.** Which TV program can catch your attention when you are channel surfing? In this paper, several models are proposed to solve the problem. Our efforts in building these models followed the path below. We establish a judgment matrix to obtain the weight of index using Analytic Hierarchy process (AHP). Then, Fuzzy Synthetic Evaluation (FSE) is built to rank the top 10 TV dramas and this can overcome weakness of excess subjective factors in AHP. Moreover, we compare our top10 with the data online and they have high similarity. A large amount of data are collected from related documents and websites. We use Grey Relational Analysis (GRA) to filter the parameters, and the conclusion is that just the number of Baidu searches and microblog followers have a direct relationship with star's popularity index. Finally, we finish a data fitting and get the index of star's popularity.

### Introduction

Compared with the film market in the United States, Domestic film market is in a bad position now. However, the development of TV show will drive and push the industry to grow rapidly and become one of the most fast-growing businesses in the future. During the "12th Five-Year Plan" period, China will vigorously develop seven key cultural industries, including film and television production. Under the circumstances of this good business environment for Chinese film and television industry, how to ensuring the quality and low the cost of the TV show become the key to receive the maximum return.

The problem needs us use mathematical modeling methods to analyze data and forecast precisely with Big Data. Based these models, we need to solve the following problems:

Question 1: Rank the TV dramas based on the ranking index and name the top 10 with the data given in Annex 1 and Annex 2.

Question 2: Actor/actress's popularity ranking could be very useful to a TV series production. We should collect and use related data as foundation, design a star popularity index, and try to prove the attainability of our index by giving a real-life example from this year.

Question 3: Based on common opinion data (such as the number of micro blogging reviews, Baidu search volume, Youku clicks, etc.) and the current audience to speculate on viewers' interest and what would make a TV show popular. We should describe an ideal production team, including the producer and actors. Moreover, we must prove our point with a real-life example.

Question 4: How to find and recommend the most interesting and suitable TV drama for each viewer based on his/her viewing history? How to find the script content most suitable for each local TV station according to the ratings of programs broadcasted by the local TV stations is also worth looking into. We will collect related data, provide a solution by using mathematical modeling methods, and use a real-life example to prove our point.

In this paper, we analyze the problems above and consult lots of literature, and then come up with the following approaches:

We build a mathematics model of ranking index based on the first-class fuzzy comprehensive evaluation (FCE), and thus several factors of TV dramas is transformed into a ranking index.

We build a model to describe the popularity index of actor/actress. The model is mainly based on the publications, age, number of Sina Microblog followers, years of debut, times for attending variety shows, number of awards and number of searches of stars. Therefore, we use the

normalization numerical of these information to calculate the index. According to the range of index, we divide the rank to describe the degree of popularity.

We collect top 50 dramas' number of micro blogging reviews, Baidu search volume, Youku clicks and their related information on the website. Then we obtain the model of popularity index, and regard them as output of Back-propagation artificial network (BP network). Then we quantity related information of dramas, including director, starring, company of production and producer using relevant method.

We collect related data of one's view history and establish a model-using cluster. Then we give the suggestions to audience and TV station by using three times clusters.

## Assumptions

We regard the TV show as TV drama. The elements that we already have taken into consideration play a vital role in the calculation of popularity index. The ignored elements of actor or TV drama (the difference of audience) do not influence the ranking. The data that we have collected is enough and accurate and the quantification is correct. There exists objective and accurate ranking for TV dramas we choose, and the rankings from selected website could reflect the accurate ranking to some extent.

## Model

Fuzzy set theory has been developed and extensively applied since 1965. It was designed to supplement the interpretation of linguistic or measured uncertainties for real-world random phenomena.

In last section, we have already articulated our metrics for ranking. Totally, there are four aspects: score, episodes, number of comments, time between date of issuance license and the first broadcast time. Before using the fuzzy set theory, we find out each TV show  $\{X_1, X_2, X_3 \dots X_{145}\}$  in each of the four aspects using the given data.

In order to determine the influence extent of each influence factors on the ranking order, we use the analytic hierarchy process (AHP) to get their weights. We transform judge matrix into weight matrix and calculate the characteristic weight  $p_{ij}$  for the  $i^{th}$  TV show's  $j^{th}$  evaluation influence factor ( $X_{ij}$ ) based on the normalized fuzzy matrix  $X_f$ . Let  $w_j$  denote the weight for each evaluation influence factor.

We use the pairwise comparison method and one-nine method to construct judging matrix  $A = (a_{ij})$ .

$$a_{ik} * a_{kj} = a_{ij} \quad (1)$$

Where  $a_{ij}$  is set according to the one-nine method.

The greatest eigenvalue of matrix A is  $\lambda_{max}$ , and the corresponding eigenvector is  $u = (u_1, u_2, u_3, \dots u_n)^T$ . Then we normalize the u by the expression:

$$x_i = \frac{u_i}{\sum_{i=0}^n u_j} \quad (2)$$

The indicator of consistency check formula:

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (3)$$

Where, n denotes the exponent number of matrix.

The expression of consistency ratio:

$$CR = \frac{CI}{RI} \quad (4)$$

A fuzzy set is defined in terms of a membership function, which maps the domain of interest, e.g. concentrations, onto the interval [0, 1]. The shape of the curves shows the membership function for each set. The membership functions represent the degree, or weighting, that the specified value

belongs to the set.

Let  $X_{ij}$  denote the  $X_j$  value for the  $i^{th}$  TV show and  $\sum_{i=0}^n u_j$  denote the sum of all the  $X_j$ . Here we use the normalization function as membership function:

$$\mu_j(X_{ij}) = \frac{X_{ij}}{\sum_{i=0}^n u_j} \quad (5)$$

After calculating  $\mu_j(X_{ij})$  for each of the  $X_{ij}$ , we could concluded the fuzzy matrix  $X_f$  ( $N$  denotes the total number of the TV show).

$$X_f = \begin{bmatrix} \mu_1(X_{11}) & \dots & \mu_1(X_{15}) \\ \dots & \mu_j(X_{ij}) & \dots \\ \mu_N(X_{N1}) & \dots & \mu_N(X_{N5}) \end{bmatrix} \quad (6)$$

So we determine the final ranking index  $Y$  for each TV show;

$$Y = W * X_f \quad (7)$$

We calculate the scores of all TV dramas given and rank the TV dramas based on the ranking index with the data given in Annex 1 and Annex 2. We and name our top 10 TV dramas as “The Annual Top Ten Series”, which are shown as follows:

Table 1 The Annual Top Ten Series

TV Drama	Score	Number of Comments	Episode	Time Interval	ranking index $Y$	Ranking
Silent Separation	6.3	26223	32	12	0.0377	1
Battle of Changsha	9.2	18419	32	42	0.0300	2
All Quiet in Peking	8.8	16084	53	14	0.0271	3
Divorce Lawyers	7.2	16872	46	73	0.0269	4
The Empress of China	5.4	13625	82	2	0.0222	5
Red Sorghum	7.5	10028	60	38	0.0188	6
Boss & Me	6.9	9285	33	8	0.0171	7
A servant and two masters	7.2	8337	43	21	0.0163	8
Honey Bee Man	7.4	7373	34	23	0.0151	9
Dating Hunter	8.3	5664	32	73	0.0135	10

We can see that the top 10 TV dramas mostly belong to modern times, and based on revolution and city. Moreover, several TV shows are directed by the same director, such as Silent Separation and Boss & Me. Therefore, the success of the TV drama depends on the subject and good team to a large extent.

We use some stars' information that will be utilized in the following models discussion as an example, and collect relative data from the Internet. We choose those 50 stars who we are familiar with, because those stars have different ages and years of debut. Thus, those stars are applicable and typical.

Searching from the Baidu Encyclopedia, a website that can provide specific data about stars, we can find relative data of our specific evaluation norms. Using this website, we finally conclude the relative statistics of those 50 stars and list them in a form. Here we give statistics of 10 stars as an example.

In the following table, “P”, “A”, “NS”, “YD”, “TA”, “NA”, and “NS”, refer to “publications”, “age”, “number of Sina Microblog followers”, “years of debut”, “times for attending variety shows”, “number of awards”, and “number of searches”, respectively.

**Table 2 . The relevant data of 10 stars**

Name	P	A	NSM	YD	TA	NA	NS
Han Lu	19	26	2576	5	50	10	6310
Long Cheng	117	52	2837	45	63	102	7770
Chao Deng	45	37	5344	14	59	57	5420
Liti Zhong	44	46	305	21	1	16	3700
Jiong He	33	42	8221	18	255	35	3250
Zhiwei Zeng	220	63	57	40	12	20	1320
XiyuanXu	26	40	3924	15	255	13	2750
Yifeng Li	20	29	3435	10	16	46	7160
Bingbing Fan	88	35	5428	18	2	105	8120
Weijia Li	4	40	2677	17	255	1	2790

Grey system theory puts forward the concept of grey correlation degree, it is the measure of the correlation between two factors in the system metrics, and correlation degree reflects the influence degree of various factors of the system directly. To use entropy method, there are mainly five steps:

Select a reference sequence as shown below:

$$x_0 = \{x_0(k) | k = 1, 2, \dots, n\} = (x_0(1), x_0(2), \dots, x_0(n)) \quad (8)$$

And the other group of sequence is,

$$x_i = \{x_i(k) | k = 1, 2, \dots, n\} = (x_i(1), x_i(2), \dots, x_i(n)), i = 1, 2, \dots, m \quad (9)$$

In this question, influence factors have different numbers of different dimension, but when we calculate the correlation degree, it requires the same numbers of same dimension. So we want to carry out a variety of data processing dimensionless. in addition, For comparison easily, all the sequences are required to have a common point. In order to solve these two problems, we transform the given sequences. Let  $x_i$  denote the  $i^{th}$  influence factor,  $x_{max}$  denote the maximum  $x_i$  value for all the stars and  $x_{min}$  denote the minimum  $x_i$  value for all the stars, we name

$$x' = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad (10)$$

Calculate the grey relational coefficient:

$$\xi_i(k) = \frac{\min_s \min_t |x_0(t) - x_s(t)| + \rho \max_s \max_t |x_0(t) - x_s(t)|}{|x_0(t) - x_s(t)| + \rho \max_s \max_t |x_0(t) - x_s(t)|} \quad (11)$$

Calculate the correlation degree of  $x_i$  to  $x_0$  is,

$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k) \quad (12)$$

We use  $r_i$  to describe the correlation degree between  $x_i$  and  $x_0$ , namely to describe the influence on  $x_0$  caused by the change of  $x_i$ .

Then, we order the correlation degrees in order of size, and we can obtain several top influence factors from this order. And we consider these factors as the most important factors of popularity index.

We can give priority to these values as the independent variables of popularity index:

$$r_1, r_2, \dots, r_l (l \leq m) \quad (13)$$

Here we use the normalization numerical as independent variables of popularity index:

$$a_i = \frac{r_i}{\sum_{i=0}^m r_i} \quad (14)$$

We can determine the final popularity index  $I$  for each star:

$$I = a_1x_1 + a_2x_2 + \dots + a_lx_l \quad (15)$$

Comparing Yang's popularity index with table of Rank Level, we can learn that Yang is extremely popular now. The results accord with the actual situation in present, so it proves that our model is reasonable and feasible.

## Reference

- [1] LONG Jiang-qi, LAN Feng-chong, CHEN Ji-qing, and YU Ping. Mechanical Properties Prediction of the Mechanical Clinching Joints Based on Genetic Algorithm and BP Neural Network [J]. Chinese Journal of Mechanical Engineering, 2009, 44(6): 27–35. (in Chinese)
- [2] TU Juan-juan. The Research on Predicting TV Audience Rating Based on the Data Mining Technology [D]. 12 March 2007.
- [3] DENG Xue, LI Jia-ming, ZENG Hao-jian, CHEN Jun-yang, ZHAO Jun-feng. Analytic Hierarchy Process (AHP) Weight Calculation Method for the Analysis and Application Research [J]. Mathematics in practice and theory, Apr, 2012, 42(7). (in Chinese)
- [4] JIANG Ming-hui, XU Pei, REN Xiao, CHE Kai. Research on algorithms development and optimization for personal credit scoring [A]. Journal of Harbin Institute of Technology, May 2015, 47(5). (in Chinese)
- [5] WU Li-yun, QU Zhou-qing. The Prediction of Audience Rating Research Based on BP Networks. Journal of Communication University of China, Sept, 2011, 18(3). (in Chinese)
- [6] LIU He. The Design and Implementation of a Recommender System Based on Clustering Algorithm [D]. 2005 (in Chinese)
- [7] LI Da-wei. The Design and Implementation of a Recommender System Based on Clustering Algorithm [D]. March 2009 (in Chinese)
- [8] LIU Ze-lin, PENG Chang-hui, XIANG Wen-hua, TIAN Da-lun, DENG Xiang-wen, ZHAO Mei-fang. Application of artificial neural networks in global climate change and ecological research: An overview. Chinese Science Bulletin, December 2010 55(34). (in Chinese)
- [9] Li Shi-xu. The Design realization of a Music Recommender System [D]. December, 2014. (in Chinese)
- [10] Tang Ming. Video Recommendation Based On User Preference. [D] June, 2013 (in Chinese).