

Application of ELO algorithm in migration of marriage matching

Tianyi Zhou

Institute of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, 222200, China

475520879@qq.com

Abstract. In this work, we mainly explore the application of the ELO algorithm in the marriage and marriage matching algorithm. With the rapid development of network communication technology, more and more people have begun to choose marriage and love on the internet, and various marriage and love applications are also emerging one after another. We will elaborate on the advantages of the ELO algorithm by comparing the ELO algorithm with the Gale Shapley algorithm mainly used in the current dating market. The high accuracy and absolute objectivity of the ELO algorithm effectively avoid the pairing generated by subjective will in the pairing of big data population, which is conducive to long-term development and communication, and also meets the requirements of ethics of both sexes from a scientific perspective. Today, with the gradual expansion of the marriage market, the simplicity of Gale Shapley's algorithm makes matching inefficient and poor. We demonstrate the comprehensiveness of the ELO algorithm in detail through mathematical methods such as hierarchical formula and finally get the interactive application in the software.

Keywords: ELO, Matching algorithm, Grading, Marriage market

1. Introduction

1.1. Introduction to the background of the ELO algorithm

The ELO algorithm is a mathematical model used to assess the abilities of competitors. Arpad ELO proposed it in the 1960s to assess the ability of chess players. However, as time passed, the application field of the ELO algorithm gradually expanded, and it is now widely used in a variety of competitive games such as e-sports, bridge, football, and baseball. In these fields, the ELO algorithm can not only provide rankings and grades but also predict game outcomes and evaluate player performance. As a result, the ELO algorithm is an important tool for assessing competitive ability. The basic idea behind the ELO algorithm is to adjust the level of players by comparing their strength gaps. This algorithm was first used in chess and has since spread to other competitive activities.

The central concept is to adjust the level of players based on the outcome of the game. Participants in the competition are assigned an initial level, which will change as the competition progresses. The algorithm's key parameters and variables include the player's initial level, the game result, and the expected winning rate.

1.2. Introduction to the background of Gale Shapley algorithm

The Gale Shapley algorithm is a classic algorithm for solving the problem of stable marriage-matching. Matching men and women is a classic combinatorial optimisation problem that involves determining how to match a group of men with a group of women so that everyone can find their ideal spouse. Because this problem was first proposed in 1962 by American mathematicians David Gale and Lloyd Shapley, it is known as the Gale Shapley algorithm.

The Gale Shapley algorithm is based on the concept of “stable matching,” which means that there is no pairing between men and women, and they prefer to be with each other rather than with the currently paired object. Iterative matching is used by the algorithm to gradually achieve a stable matching result.

Through the algorithm’s core steps, everyone will make a choice based on their own preferences, and each iteration will force those who are not selected to propose again. This method of pairing and alternating not only ensures that everyone finds a suitable spouse but it also ensures that the final matching result is stable and there is no better pairing.

2. Algorithm mechanism of ELO

2.1. Performance grading formula

An important formula in the ELO algorithm is as follows:

$$Rp = Rc + D(p) \quad (1)$$

The performance rating is marked as Rp , and the opponent’s ELO score is marked as Rc . When there are multiple opponents, the average ELO score is calculated. $D(p)$ is the inverse function of $P(D)$. That is, the average winning rate is used to estimate the equal scored in turn.

For example, an inexperienced player and a 1700-point chess player played 10 games, winning 9 games and losing 1 game, so his graded ELO performance score is calculated as follows:

$Rc=1700$, $p=0.9$, and reverse $D(p)=366$ (using the score difference and expected victory table), then $Rp=1700+366=2066$, that is, this player is equal to the player who scored in 2066, and his real rating is close to 2066.[1]

2.2. Current performance sub-formula

The performance grading formula is primarily used for quickly grading or evaluating players, but when the timeline is extended, it becomes difficult to reflect a player’s true strength. Because a player’s true strength gradually improves, his later performance will outperform his earlier performance. A more accurate method is to synthesise the performance grades to calculate the players’ current strength, which is known as the ELO integral. The ELO integral is constantly changing, but in general, it will converge with the actual strength of the players.

In order to approach the actual level of player A, there are three steps:

1. Calculate the expected winning rate of A according to the difference between the two sides.
2. A and B compete to get the result.
3. Update the scores of A and B according to the outcome.

The specific formula is as follows: [2]

$$Rn = Ro + K * (G - Ge) \quad (2)$$

Rn is the new ranking score after the game;

Ro is the ranking score before the game;

K is the weight coefficient. Generally speaking, the higher the level of competition, the smaller the value of K .

(When the score is < 2000 , $K=30$; $2000\sim2400$, $K=20$; >2400 , $K=10$)

G is the player’s actual score (1 for winning, 0 for losing and 0.5 for drawing).

Ge is the player’s expected score (expected winning percentage * 1 + expected draw probability * 0.5) based on the original ranking.

3. Migration application and improvement of ELO algorithm

3.1. *ELO migration application in marriage and love software*

ELO algorithm has the advantages of good matching, stable interval planning and high timeliness. Under the accumulation of big data, the final evaluation effect will be significantly better than short-term data matching, and it can effectively reflect the level differences of individuals in the evaluation software with complicated data and long time.

3.1.1. Matching Mechanism of ELO in Marriage Software. The ELO algorithm must be applied to a massive data matching system, and personal evaluation can be obtained by using the public's general evaluation. As a result, the software of marriage and love is overly reliant on the subjective will to rank one-on-one through personal wishes and then use a ranking algorithm for instant matching. The ELO algorithm can effectively divide people into groups and use the integral system to objectively score each user based on other ratings. The ELO algorithm employs the "high score, low score, and low rise" strategy. When users register the software, they will receive a basic score. Each user will begin at the same point, and during the course of using the software, the score will be "marked as interested" or "marked as uninterested" by others in order to calculate the rating. When users are praised and followed by high-scoring users, they gain a lot of points, but they do not gain a lot of points when they are praised and followed by low-scoring users.

We may believe that being praised and paid attention by low-scoring users is natural, but getting the attention of high-scoring users is relatively difficult. Add and subtract scores by combining the difference between both parties' scores using the following calculation formula:

$$R_n = R_o + K * (G - G_e) \quad (3)$$

When a user's score tends to be stable for a long time, it can be considered that the user is in an exact interval. At this time, the software can recommend the opposite sex in the same interval to the user accordingly to make the matching stable.

In the whole process, each user's ELO score is a hidden score collected by the system background, and users themselves can't know their own scores.

3.1.2. Advantages of ELO compared with Gale Sharpley algorithm. The ELO algorithm can systematically and objectively assess users' level of love and marriage, whereas the Gale Shapley algorithm is primarily dedicated to fast matching, and users' subjective wishes can be used for stable and fast optimal matching. The ELO algorithm disregards any user's subjective viewpoint, and different users may overestimate or underestimate their own evaluation. However, the Gale Shapley algorithm's referenced factors are too few and skewed towards men. The ELO algorithm considers both parties' perspectives. Users in the same interval become acquainted with each other's hobbies and personalities through long-term screening. Furthermore, when a user is stable in an interval, the overall level of the opposite sex in the same interval effectively improves the user's previous misunderstanding of self-level, allowing both parties to recognise themselves clearly and treat each other more equally in the communication process.

To summarise, Gale Shapley pursues speed and a high success rate of matching, but this does not guarantee the success rate of two people's love. Every user can successfully match, but the chances of being called a partner are extremely low. The ELO algorithm takes a long time to investigate, but once paired, it is very likely to become a partner. In terms of each user, the ELO algorithm can effectively find the right person, whereas the Gale Shapley algorithm simply solves the problem of leftover men and women and does not consider the two people's happy lives after pairing.

3.2. Interactive application of ELO and Gale Sharpley algorithm

ELO algorithm and Gale Shapley algorithm have their own advantages and disadvantages, so we can learn from each other's strong points and make interactive applications. The specific application steps are as follows:

When users register, give priority to the investigation of users' marriage wishes. If the user has a strong desire to get married and does not require too much from his partner, we attribute the data of such users to Gale Shapley.

Database of algorithms. If the user wants to be screened for a long time and is not in a hurry to get married, ELO algorithm is used to evaluate the user, but it is not included in Gale Shapley's database. At the same time, it needs to be clear that every user needs to be included in the ELO database, because it is possible that an old user who has been screened for a long time will be fascinated by the type of a new user who has a strong desire to get married. At this time, it is necessary to weigh the priority of Gale Shapley and ELO algorithms and the subjective wishes of both parties, quantify the Gale Shapley algorithm into integral, participate in the calculation system of ELO, and then recommend it to both parties.

In addition, we can investigate the old users. If we have a strong desire to get married after using the software for 2~3 years, we can use Gale Shapley algorithm to effectively match it, and the success rate of this interactive application will be much higher.

3.3. ELO algorithm improvement

The conventional ELO algorithm is a one-to-one integral system. When the number of users of dating software is gradually expanding, there are often many people who like and pay attention to one person at the same time. At this time, an algorithm is needed, a one-to-many integration system.

$$R_n = R_o + K * (G - G_e) \quad (4)$$

G_e is the expected score of the player based on the original ranking.

(Expected Tagged Probability-Expected Rejected Probability)

For example, if a man with a score of 1600 is praised by a woman with a score of 1700 (recorded as winning), then the player's new ELO score is calculated as follows:

$R_o=1600$, $K=300$, $D=100$, and the expected winning rate is 36% according to the formula of ELO algorithm.

(Look up the table by using the score difference and expected winning rate table)

Assuming $K=30$, then

$R_n=1600+30*(1-0.36)=1619.2$, which means an increase of 19.2 points.

When many people pay attention to it, the G_e of this formula adopts the arithmetic average viewpoint, and it is similar to G .

Calculate the differential. However, because the weight of each person who likes and pays attention is not always the same, using the cumulative density method improves calculation accuracy.[3]

$$\mu_{winner} = \mu_{winner} + \frac{\sigma_{winner}^2}{c} * \varphi\left(\frac{t}{c}\right) \quad (5)$$

$$\sigma_{winner} = \sigma_{winner} * \left[1 - \frac{\sigma_{winner}^2}{c} * \omega\left(\frac{t}{c}\right)\right] \quad (6)$$

T , c , $\varphi(t)$ meaning as follows:

$$t = \mu_{winner} - \mu_{marksman} \quad (7)$$

$$c^2 = 2 + \beta^2 \sigma_{winner}^2 + \sigma_{marksman}^2 \quad (8)$$

$$\varphi(t) = \frac{N(t)}{\tau(t)} \quad (9)$$

$N(t)$ stands for probability density function, $\tau(t)$ stands for cumulative density function and β^2 stands for the degree of instability played by players. This algorithm will be more complicated in calculation, but the accuracy of calculating the difference will be higher in the case of one-to-many

4. Conclusion

4.1. Summary

Due to time and ability constraints, this paper does not provide a detailed explanation of the ELO algorithm on the losing side. Simultaneously, the rules of dating software differ slightly from those of chess. There will be a tie score rule in chess competitions, but there are only two options in dating software: marking or not marking. However, the precision of the ELO algorithm will be higher than that of the Gale Shapley algorithm, which demonstrates the transcendental assumption in statistics, reduces error through repeated experiments, and eventually approaches the real level of users, which is similar to the machine learning algorithm.

4.2. Outlook

In terms of dating software, future work can primarily focus on creating the “table of points difference and expected winning rate” and perfecting the integral calculation of expected winning rate and unmarked reporters. By addressing these two issues, the user experience will be improved more effectively, and it will also have commercial value.

References

- [1] Peng Wenzhi. China's online game industry status and development policy analysis [J]. China Collective 2008(24):148-149.
- [2] Ouyang Meisheng, Xia Hongxia, Research on Evaluation System of Competitive Events Based on ELO Algorithm [D]. Wuhan: Computer College of Wuhan University of Technology 2013:1~12
- [3] Lin Aoliu, MOBA game balance research [D]. Yunnan: Computer College of Yunnan University 2015:18~22