Analyzing Temporal Trends and User Modeling on an Online Chess Community

IN4252 - Web Science and Engineering

Santhos Baala Ramalingam Santhanakrishnan
TU Delft - Student Number 4740270
S.B.RamalingamSanthanakrishnan@student.tudelft.nl

ABSTRACT

Lichess is a crowdfunded online chess community with over 250 million recorded games since 2013. In this draft paper, we explore the public game dataset for general and temporal trends, corresponding to player levels, identify choices of top chess players over feasible time periods and answer questions that arise along the way. We also compare the online trends against the real-world ELO distribution to gain insights on specifics of the Glicko rating system, compared to the real world FIDE ELO and its effect on user retention. Using the insights we try to draft the features of an online chess player, that acts as a basis for further user modeling and merging with profiles from other social networks. From a domain point of view, we identify user preferences such as opening choices, preferred time control, frequency of visits to gauge overall involvement in the community.

1 INTRODUCTION

Lichess [1] is a free online chess community, started in 2013 and it peaked in 2017 with over a million games played every single day across various time controls and chess variants. It is crowdfunded, open source and each game is recorded, resulting in a large public dataset with rich meta data and domain-specific chess information. Chess is a cognitive-heavy game and has been time and again a sample space for understanding human psychology and the vastness in available data has been a paradise for statisticians to elicit interesting patterns, the extent to which, display of skills in chess has been correlated to excellence in academics and other cognitive abilities.

Insights gained from analyzing Lichess shall help form a basis for future work on how to model and incentivise a wide range of online learning processes in order to create strong retention, where the cognitive abilities are being challenged. Also, skill building on Lichess itself is a non-trivial and a long term problem and we need more insights on providing a tangible improvement path for users, more in the line of MOOCs [2], and recommend courses based on observed skill and behaviour.

TODO: A paragraph on the structure of the rest of the paper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

2 RELATED WORK AND DATASET

Analysis from a cognitive science standpoint has been carried out on chess databases by [2], in which the authors have taken up the German national chess database and laid out methods to carry out analysis on skill development over lifetime, birth cohort effects, effects of activity and inactivity, etc. Contrary to the analysis in [2], we do not have any skill or recording restrictions in Lichess, however lack of demographic information in the available dataset restricts the kind of analysis that we can perform.

A common feature of competitive sports is measuring the relative skill of a player, usually quantified as ELO points through a rating system (compared by [3]), of which the Glicko-2 [1] rating system is adopted by Lichess. An important difference to be noted in our dataset is that ratings of a player are calculated and assigned in real-time (at the end of each game) whereas in the real world rating lists are published periodically, typically monthly or quarterly, which means a recorded game could actually be reflecting a shadow rating. This leads to high volatility of the Glicko's *RD* factor in Lichess. For feasibility reasons, we analyze only a subset of the available data [3], between 2013-Jan-1 and 2014-Jan-31 in three game formats – bullet, blitz, classical and correspondence. The dataset, after cleanup of data, contained a total of 4,071,927 games, consisting of bullet - 1,086,961 (26.694%), blitz - 1,576,431 (38.715%), classical - 1,399,978 (34.381%) and correspondence - 8,557 (0.21%) games respectively. A total of 54,058 unique players and their corresponding profiles were captured from the dataset. The list of features of a game and a player are given in detail along with an example in appendix.

3 ANALYSIS

We first start with basic temporal observations to develop an understanding of the dataset and then move on to features that we consider for user modeling. Figure 1 shows the timeline of games played in various formats, we observe that there are spikes and valleys, which were found to be strongly correlating to new users (coefficient of 0.87) joining daily [4], no relation was found between days of week though.

[1] http://lichess.org
[2] Lichess TV: https://lichess.org/tv, Lichess Learn: https://lichess.org/learn
[3] http://database.lichess.org
[4] One caveat is that, day refers to an UTC day throughout the paper.

Figure 1: Games played over time by type

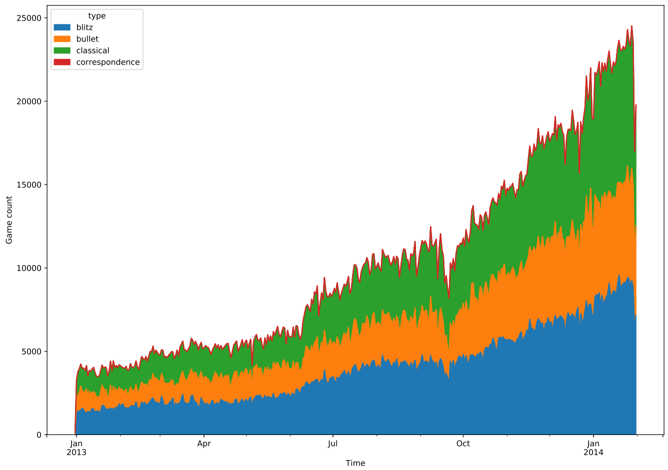

Figure 2 shows the probability density function of player ratings (elo) in the system, by min, max and mean ratings. We clearly see that there is huge spike at 1500 and sub-peaks at around 700 points above and below 1500. This is contrary to the analysis in [2] and FIDE elo ratings in general. This is due to the fact that real-world rating systems never assign a rating to a new player unlike Lichess, which assigns 1500, based on the Glicko-2 documentation [1]. A close correlation was found between elo points and the *RD* factor (-0.007658) and due to skew from a large number of data points at 1500, it was slightly negative stabilizing as it approached higher elos, visible through descriptive statistics in Table 1. The max and min *RD* explains the sub-peaks in Figure 2. This leads us to the question *Are weaker players majority contributors to the Lichess database?*.

Figure 2: Rating distribution in blitz games

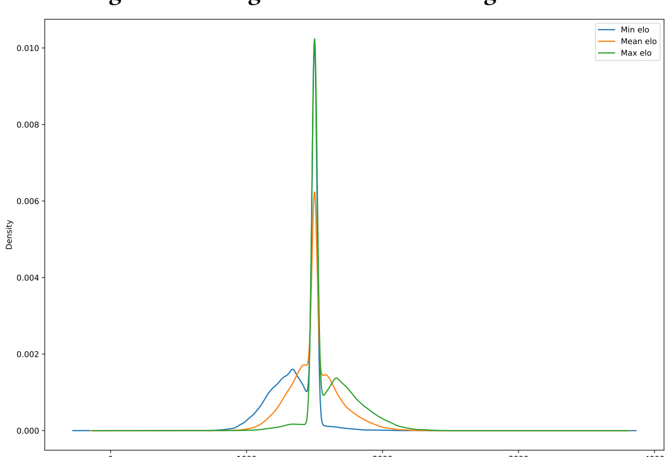

Table 1: Elo and *RD* factor descriptive statistics

Stat	Elo	RD
mean	1605.25	-0.671
std	214.83	34.64
min	0	-624.000
25%	1467.00	-10.00
50%	1607.00	0.00
75%	1751.00	10.00
max	2828.00	649.00

4 DISCUSSION

5 CONCLUSION

REFERENCES

[1] M. E. Glickman. The glicko system. *Boston University*, 1995.
[2] N. Vaci and M. Bilalić. Chess databases as a research vehicle in psychology: Modeling large data. *Behavior research methods*, 49(4):1227–1240, 2017.
[3] N. Veček, M. Črepinšek, M. Mernik, and D. Hrnčič. A comparison between different chess rating systems for ranking evolutionary algorithms. In *Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on*, pages 511–518. IEEE, 2014.

APPENDIX