ELO Rating System in Practice

Jeremy Madison

Formulas copied from “Elo rating system” on Wikipedia
Problem

- Large body of players
- All can't play each other
  - Ex: 10000 players playing 250 games/year
  - Requires 40 years to finish
- How to estimate rank?
Model

- Tournament is weighted digraph $G = (V, E)$
- Each $v$ in $V$ is a player
- Each $e$ in $E$ indicates two players played
- Edge weight is expected result formula from player $v_i$'s perspective
- Rating change formula:

$$R'_A = R_A + K(S_A - E_A).$$
More Formulas

• Expected Outcome:

\[ E_A = \frac{1}{1 + 10(R_B - R_A)/400} \]

• K formula:

\[ K = \frac{800}{(N_e + m)} \]

• K usually assumed to be 32

• \( S_a \) is actual outcome
Outcome Possibilities

• For the coming example:
  – 3 possible values for $S_a$
  – 0 for a loss
  – .5 for a draw
  – 1 for a win
Single Game Example

- Player A is rated 1600
- Player B is rated 2000
- A's expected result:
  \[
  1/(1+10^{[2000 - 1600]/400}) = 1/11
  \]
- B's expected: 10/11
- Player A miraculously wins!
- New Rating: 1600 + 32(1 – (1/11)) = 1629
- B Rating: 2000 + 32(0 – (10/11)) = 1971
Example Tournament

• Hypothetical 5-round event
• Player A's (rated 1600) results:
  • Wins: 2000, 1400
  • Draws: 1600
  • Losses: 1700, 1300
Partial Tournament Graph

- Vertices = Players
- Edges = Game Played
- Direction = Result
- Weight = Expected Result
Results for Player A

- \( S_a = \sum \text{Actual Results rounds 1-5} = 2.5 = \text{out}(A) + 0.5 \times \text{nodir}(A) + 0 \times \text{in}(A) \)
- \( E_a = \sum \text{Expected Results rounds 1-5} = 2.559610... = \sum \text{edgeweight}\{A,i\}, \text{ for all } i \text{ opponents of } A \)
- \( K = 32 \)
- Change: \( 32(2.5 - 2.55...) \)
- New Rating for Player A: 1598
Newcomer Problem

• New, unrated player
• How to get a rating?
• For first tournament, use Performance Rating
• Performance: Average of all round results
  – Win = Opponent Rating + 400
  – Loss = Opponent Rating – 400
  – Draw = Opponent Rating
• From previous example: Perf(A) = 1600
Conclusions

• The ELO rating system is reliable for a large body of players
• Ratings tend towards actual skill over time
• Can accommodate players entering and leaving system
Questions

Any Questions?
Sources

Batchelder, William H.; Bershad, Neil J.


Slutzki, Giora; Volij, Oscar.

Stefani, Ray; Pollard, Richard.