

Průvodce pro začátečníky k revidovanému systému hodnocení ICCF

Mark Glickman, Katedra statistiky, Harvardská Univerzita

Počínaje rokem 2023 přijme ICCF nový systém hodnocení pro hodnocení výsledků dosažených v soutěžích ICCF. Rozhodnutí revidovat ratingový systém bylo výsledkem zjištění, že starý systém již není schopen adekvátně vypočítat přesné ratingy. To bylo zvláště patrné na nejvyšších úrovních hry, kde frekvence partií ukončených remízou vzrostla do bodu, kdy se hodnocení vysoko hodnocených hráčů téměř nezměnilo.

V roce 2021 delegáti ICCF uznali potřebu nového systému, který by řešil problém zvýšené frekvence remízových partií. Za tímto účelem zadali vývoj systému na míru Dr. Marku Glickmanovi, uznávanému expertovi na ratingové systémy a docentovi na Harvardské univerzitě. Dr. Glickman je znám svou prací na systémech hodnocení Glicko a Glicko-2, které byly implementovány v různých šachových organizacích, jako jsou [chess.com](https://www.chess.com) a lichess.org. Byl také jedním ze spoluvývojářů univerzálního systému hodnocení používaného na Grand Chess Tour.

Dr. Glickman vyvíjel revidovaný systém ICCF od srpna 2021 do května 2022 s použitím více než šestiletých výsledků partií ICCF a aktuálních informací o ratingu k doladění vzorců. Podrobnosti nového systému byly představeny delegátům ICCF v červnu 2022 a přijaty byly v srpnu 2022. Implementace vzorců pro použití ICCF začala krátce poté.

Tento dokument popisuje základní vlastnosti nového ratingového systému a je určen pro netechnické publikum.

Rating a odchylky ratingu:

Jednou z významných změn starého ratingového systému je přidání odchylky ratingu (zkráceně RD) pro každého hráče. Jedná se o koncept, který byl vypůjčen z ratingového systému Glicko vyvinutého zhruba před 30 lety. Rating, stejně jako ve starém systému

ICCF, je měřítkem hráčské síly. RD je měřítkem nespolehlivosti ratingu. Čím větší je RD hráče, tím méně spolehlivé je hodnocení hráče. RD kolem 75 nebo nižší znamená, že hodnocení hráče je spolehlivé a důvěryhodné.

Hráči, kteří nemají rating (za předpokladu neexistence žádného jiného externího ratingu, jako je rating FIDE), začínají s ratingem 1800 a RD 250, což je největší možná RD v novém systému. Toto přiřazení znamená, že náš nejlepší odhad ratingu hráče bez jakýchkoli dalších informací je 1800. Nicméně RD 250 znamená, že hodnocení 1800 je velmi nejisté. Pro hráče s ratingem FIDE, kteří nemají rating ICCF, je počáteční RD nastaveno na 150. Hráči, kteří mají vysoké RD, jsou obvykle buď prozatímní hráči, nebo ti, kteří v poslední době nesoutěžili a mají zastaralý rating. Naopak hráči, kteří mají zavedené ratingy, a kteří často soutěží, mívají nízké RD. Začlenění RD do revidovaného ratingového systému ICCF má některé důležité důsledky pro hodnocení partií. RD mohou ovlivnit, jak moc se změní hodnocení hráče od výsledků partie a také jak moc se změní hodnocení jeho soupeře.

Řekněme například, že dva hráči, Sára a Josef, se stejným ratingem, hrají partii a Sára vyhraje. Pokud mají oba hráči nízkou RD, což znamená, že jejich ratingy jsou spolehlivým odrazem jejich herní síly, pak nový ratingový systém pravděpodobně příliš nezmění ratingy Sáry a Josefa, protože jejich ratingy byly spolehlivé již před partií. Ale pokud by Sářina RD byla nízká a Josefova RD velká, což naznačuje, že jeho rating je nespolehlivý, pak by se Sářin rating stěží zvýšil, protože porazila soupeře s nespolehlivým ratingem. Josefův rating by se však pravděpodobně výrazně snížil, protože prohrál se soupeřem s přesným ratingem a jeho vlastní rating byl zpočátku nespolehlivý. A konečně, pokud by Sářina RD byla vysoká a Josefova RD byla nízká, pak by Sářin rating výrazně vzrostl, protože její rating byl nespolehlivý a Josefův přesný.

Revidovaný ratingový systém ICCF byl navržen tak, aby hráčům s vysokou RD poskytl příležitost zlepšit svůj rating rychleji než dříve. Protože hráči s vysokou RD mají tendenci dosahovat vyšší zvýšení ratingu (v absolutních číslech) než hráči s nízkou RD, hráč s vysokou RD může hrát více partií a rychle si zlepšit svůj rating v relativně krátkém čase. Tento přístup má za cíl motivovat hráče, kteří jsou noví v soutěžích ICCF, aby začali hrát

partie zahrnuté do ratingu, a odměňovat prozatímní hráče, kteří soustavně soutěží a účastní se více partií zahrnutých do ratingu.

Postup stanovení ratingu:

Revidovaný ratingový systém ICCF aktualizuje rating hráčů a RD každé tři měsíce, stejně jako starý systém. Zde jsou kroky zahrnuté v procesu hodnocení, který se opakuje každé čtvrtletí:

1. Určete rating a RD každého hráče na začátku 3měsíčního ratingového období.
2. Aktualizujte rating a RD každého hráče pomocí nových vzorců ratingu na základě partií dokončených během 3měsíčního období. Tento krok bude mít za následek snížení RD pro každého hráče.
3. Použijte vzorce ke zvýšení RD každého hráče kvůli plynutí času na vytvoření nové RD, který se použije na začátku příštího 3měsíčního období.

Pro upřesnění tohoto procesu stojí za zmínku několik poznámek. Za prvé, na začátku každého ratingového období je každému hráči přidělen rating a RD, i když nemají rating. Hráčům bez ratingu je přiřazen rating a RD (obvykle vysoká hodnota), které nejsou založeny na výsledcích partií ICCF. Mohly by být založeny na známých ratingových informacích z jiného ratingového systému (např. FIDE), nebo, jak je uvedeno výše, by jim mohly být přiděleny rating 1800 a RD 250, pokud nejsou k dispozici žádné další informace.

Za druhé, všechny výsledky partií za 3 měsíce jsou hodnoceny současně. To znamená, že hodnocení a RD každého hráče jsou aktualizovány současně na základě všech partií dokončených během ratingového období. Zatímco rating hráče může růst nebo klesat, RD na základě kroku 2 vždy klesá. To odráží myšlenku, že jak je zaznamenáno více výsledků partií, stává se aktualizovaný rating hráče spolehlivějším měřítkem jeho herní síly.

Nakonec je důležité poznamenat, že po snížení RD kvůli výsledkům partií v kroku 2 se RD následně zvýší v kroku 3, aby se zohlednilo plynutí času, a tato nová hodnota se

použije na začátku příštího 3-měsíčního ratingového období. I když hráč během období nedokončí žádnou partii, jeho RD se přesto zvýší. Je to proto, že jak čas plyne, existuje více nejistoty ohledně hráčských schopností – mohli by aktivně zlepšovat svou hru, nebo by mohli zrezivět zanedbáváním šachů. Nárůst RD odráží myšlenku, že si po několika měsících jsme o něco méně jisti silou hráče. Je však třeba poznamenat, že RD nad 120 se nezvyšuje v důsledku plynutí času; mohou se zvýšit pouze tehdy, když RD klesnou pod 120.

Separátní pravděpodobnosti výher, proher a remíz:

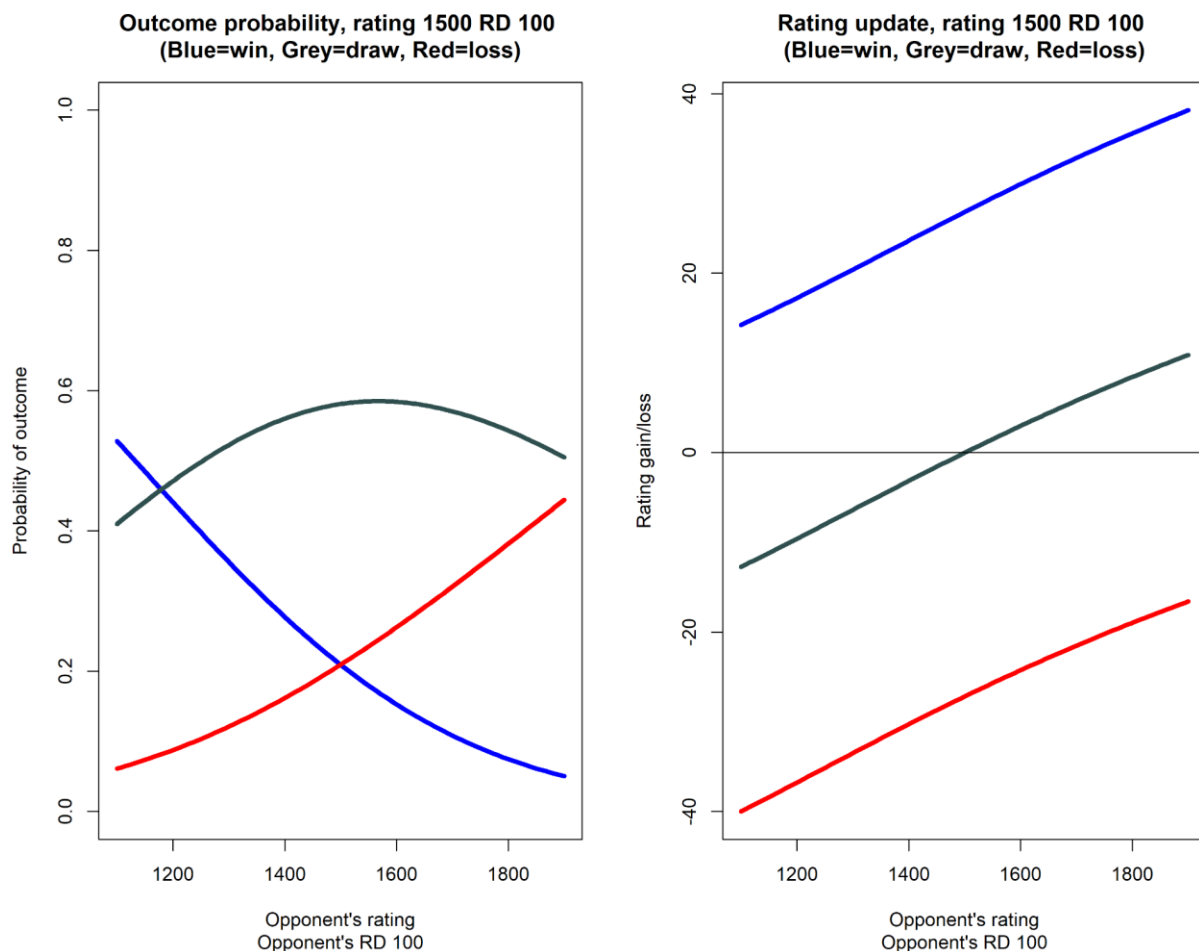
Jednou z hlavních motivací pro revizi systému hodnocení ICCF byla podstatně vyšší frekvence remízových partií mezi hráči s nejvyšším ratingem ve srovnání s hráči s nižším ratingem. To vedlo k typicky menším změnám ratingu u špičkových hráčů, kteří zřídka mají rezultativní partie proti soupeřům podobné síly. Jedním z omezení předchozího systému bylo, že byl založen na vzorcích Arpada Ela z 50. let, které nepočítaly s pravděpodobností remízy na základě úrovní schopností hráčů. Systém se spoléhal pouze na vzorec k určení očekávaného výsledku nebo očekávání výhry v partii, který byl založen čistě na rozdílu ratingu mezi dvěma hráči. Výsledkem bylo, že systém nebyl schopen vyřešit problém vysoké frekvence remízových partií mezi hráči s nejvyšším ratingem.

Revidovaný ratingový systém ICCF obsahuje tři separátní pravděpodobnosti pro každý typ výsledku partie (výhra, prohra nebo remíza), namísto jediné naděje na výhru založenou pouze na rozdílu v ratingu mezi dvěma zúčastněnými hráči. Pravděpodobnosti každého výsledku se počítají na základě ratingu a RD hráčů. Čím blíže jsou dva ratingy u sebe, tím vyšší je pravděpodobnost remízy. Čím vyšší jsou RD pro dva hráče, což ukazuje na nespolehlivost hodnocení hráčů, tím vyšší je pravděpodobnost remízy. Revidovaný systém navíc uznává, že hráči s vysokým ratingem mají tendenci spolu remízovat rychleji než hráči s nižším ratingem. Důsledkem uznání, že hráči s vysokým ratingem spolu tahají vysokou rychlostí, je, že zisky a ztráty ratingu založené na rezultativních partiích mohou být výraznější. Například, když hráč s ratingem 1500 porazí hráče s hodnocením 1300, získá rating 20 bodů (za předpokladu RD 100 pro oba hráče),

zatímco u hráče s ratingem 2500, který porazí hráče s hodnocením 2300, se rating zvýší o 24 bodů, což je více než zisk z ratingu pro dvojici hráčů s nižším ratingem.

Vzorce pro výpočet pravděpodobností každého výsledku partie a vzorce pro změny ratingu založené na těchto pravděpodobnostech byly odvozeny z analýzy předchozích výsledků partií v ICCF za šest let. Optimalizace ratingového systému s jediným cílem vytvořit přesné pravděpodobnosti by způsobila změnu ratingu o příliš velké hodnoty a ratingy nejlepších hráčů by se stěží změnily na základě remízových partií. Místo toho bylo odvození vzorců kompromisem mezi třemi faktory: (1) vytvářením přesných výpočtů pravděpodobnosti výsledků partií, (2) zabráněním inflace ratingu a (3) zachováním rozložení ratingu v průběhu času. Zatímco míra remízových partií mezi top hráči v posledních letech přesahuje 95 %, konečné vzorce počítají s pravděpodobností kolem 80 %. Tato mírná nepřesnost umožňuje větší pohyb ratingu na nejvyšších úrovních a zabraňuje nechtěným posunům v průměru všech ratingů.

Níže jsou grafy, které ilustrují pravděpodobnosti výhry, remízy a prohry pro hráče s určitým ratingem a RD, a odpovídající změnu ratingu pro každý výsledek partie.

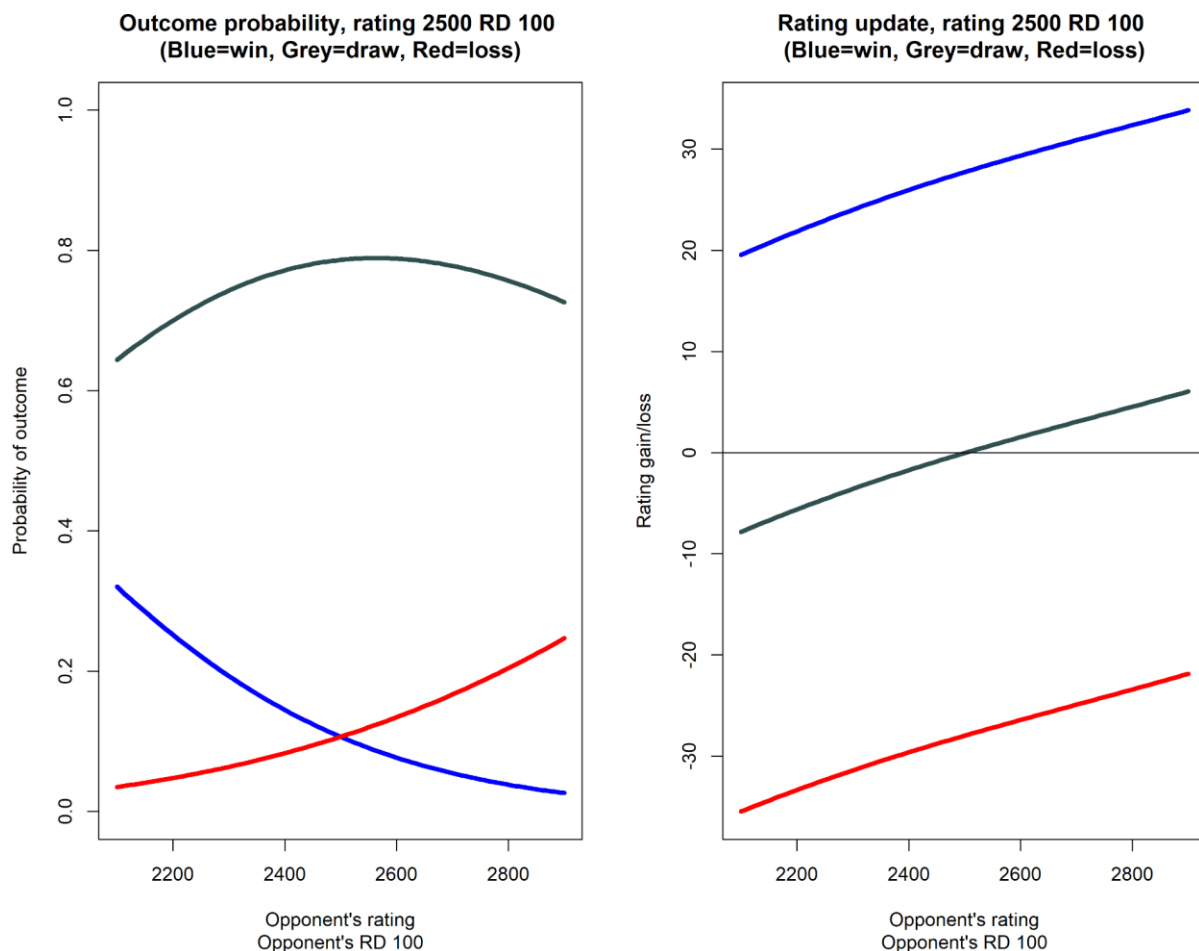


Obrázek výše odpovídá hráči s hodnocením 1500 a RD 100. Graf vlevo ukazuje pravděpodobnost výhry (modrá křivka), remízy (šedá křivka) a prohry (červená křivka) proti soupeři s ratingem vyznačeno na vodorovné ose (a s RD 100). Například, když hrajete proti soupeři s ratingem 1400 a RD 100, pravděpodobnost prohry s tímto soupeřem je asi 0,16, pravděpodobnost remízy je asi 0,565 a pravděpodobnost výhry je asi 0,28. Pravděpodobnost remízy je největší (nejvyšší bod na šedé křivce v levém grafu), když je rating soupeře také 1500. Za zmínku stojí, že pravděpodobnost remízy proti soupeři s ratingem 1900 (o 400 bodů vyšší než 1500- hodnocený hráč) je o něco vyšší než pravděpodobnost remízy proti soupeři s hodnocením 1100 (o 400 bodů méně než hráč s ratingem 1500), jak lze vidět porovnáním výšky šedé křivky na pravé a levé straně grafu. Důvodem je, že revidovaný ratingový systém uznává, že dvojice hráčů s vyšším ratingem mají tendenci remízovat partie častěji než dvojice hráčů s nižším ratingem. V tomto případě je pravděpodobnost remízy mezi hráči s hodnocením 1500 a 1900 vyšší

než pravděpodobnost remízy mezi hráči s hodnocením 1500 a 1100, protože první soubor ratingů je v průměru vyšší než druhý. 1500 a 1100, protože první soubor ratingů je v průměru vyšší než druhý.

Graf vpravo ukazuje změnu ratingu na základě porážky (modrá), remízy (šedá) a prohry se (červená) soupeřem s ratingem uvedeným na vodorovné ose. Pokud je rating soupeře 1500 (a jeho RD je 100), rating hráče se nemění, pokud je výsledkem remíza, ale získá ratingu je asi 27 bodů za výhru a pokles ratingu o 27 bodů za prohru. Opět stojí za zmínku asymetrie ve změně ratingu při partii se soupeřem s ratingem o 400 bodů vyšším versus o 400 bodů nižším. Například při remíze se soupeřem s ratingem vyšším o 400 bodů je získá ratingu pro hráče s ratingem 1500 11 bodů. Avšak při remíze se soupeřem s ratingem o 400 bodů nižším je ztráta ratingu pro hráče s ratingem 1500 13 bodů. Tato asymetrie odráží, že remíza se soupeřem s vyšším ratingem je pravděpodobnější než remíza se soupeřem s nižším hodnocením, protože šance na rezultatívni výsledek je vyšší.

Stejná analýza může být ilustrována na hráči s ratingem 2500 a RD 100.



Ve srovnání s analýzou hráče s ratingem 1500 je třeba poukázat na několik rozdílů. Na základě grafu vlevo je pravděpodobnost remízy výrazně vyšší než 0,6, když hráč s ratingem 2500 hraje proti soupeři s ratingem mezi 2100 a 2900 (a RD 100), s vrcholem přibližně 0,8, když rating soupeře je také 2500. To je mnohem vyšší než pravděpodobnost remízy hráče s ratingem 1500, který hraje proti soupeři v rozsahu 400 ratingových bodů. Graf vpravo zobrazující změnu ratingu v důsledku výhry, remízy nebo prohry má některé důležité vlastnosti, které stojí za zmínku. Zatímco zvýšení ratingu za výhru proti soupeři se stejným ratingem je přibližně stejné (27 bodů pro hráče s ratingem 1500 a 28 bodů pro hráče s ratingem 2500), zvýšení ratingu za výhru hráče s ratingem 2500 nezávisí tolik na ratingu soupeře jako u hráče s ratingem 1500. Jinými slovy, hráč s ratingem 2500, který porazí soupeře s ratingem o 200 bodů nižším (s RD 100), má za následek zvýšení ratingu o 24 bodů, zatímco hráč s ratingem 1500, který porazí soupeře s ratingem o 200 bodů nižším, získá bodový nárůst pouze 20 bodů. Křivky na pravém grafu pro hráče 2500 jsou

„plošší“ než křivky na pravém grafu pro hráče s hodnocením 1500. To znamená, že pro hráče s vyšším ratingem mohou rezultativní výsledky partií vést k výraznějším změnám ratingu.

Závěr:

Revidovaný ratingový systém uznává, že dvojice hráčů s vyšším ratingem mají tendenci remizovat partie častěji než dvojice hráčů s nižším ratingem, což má za následek poněkud větší změny ratingu pro rezultativní výsledky partií u hráčů s vyšším ratingem. Zařazení RD jako míry nespolehlivosti ratingu navíc přináší výhody hráčům všech ratingů, což vede k přesnějším hodnocením, která lépe odrážejí úroveň dovedností hráče. Pro přechod na nový ratingový systém použije ICCF jako výchozí bod stávající hodnocení z doby před několika lety. Po spuštění nového systému se spustí RD hráče na základě počtu nedávno dokončených partií. Další podrobnosti o zavedení nového systému jsou k dispozici. Hráči, kteří se chtějí dozvědět více o technických podrobnostech revidovaného algoritmu hodnocení, si mohou přečíst technické specifikace zde:

https://webfiles.iccf.com/rules/Rating_formulae_for_ICCF_system.pdf.

On-line kalkulačka, která implementuje nové vzorce hodnocení a umožňuje hráčům vypočítat si vlastní změny ratingu: <https://ratingscalculator.azurewebsites.net/>.

(Překlad z anglického originálu: Ing. Josef Mrkvička)