

Procena mapiranja šahovskih rejtinga na više platformi sa modalnom regresijom

ZKSKPROT1ZKSK Istraživačke beleške

29. maja 2026. godine

Abstract

Igrači često pitaju kako se ocene na ZKSKPROT0ZKSK upoređuju sa ocenama na ZKSKPROT2ZKSK. Pitanje je praktično važno za ZKSKPROT1ZKSK jer je njegovo merilo krive se grade od veoma velikog slojevitog uzorka uzetog iz kompletnog ZKSKPROT0ZKSK mesečne istorije igara, dok korisnici mogu donositi istorije iz bilo koje sajtu. Ova napomena opisuje trenutni metod koji se koristi za procenu mapiranja specifična za tip igre od ocena ZKSKPROT3ZKSK u ocene ZKSKPROT2ZKSK. The metoda koristi podudaranja kandidata sa istim korisničkim imenom, tačnu ili skoro tačnu kontrolu vremena filteri, minimalni zahtevi za istoriju igre i procedura modalne regresije dizajniran da smanji pristrasnost od lažnih podudaranja istog korisničkog imena. Struja mapiranje proizvodnje je:

$$\hat{R}_{\text{Chess.com}} = \alpha + \beta R_{\text{Lichess}}.$$

Procene su preliminarne i biće ažurirane kako odgovarajući uzorak raste, posebno za brze igre duže od 10 minuta.

Trenutno mapiranje proizvodnje

Za čitaoce koji prvo žele rezultat, trenutne proizvodne jednačine su:

$$\begin{aligned}\hat{R}_{\text{Chess.com,bullet}} &= -530.60 + 0.9859R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,blitz}} &= -548.88 + 1.0837R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,10minrapid}} &= -499.68 + 1.0778R_{\text{Lichess}}, \\ \hat{R}_{\text{Chess.com,>10rapid}} &= -351.72 + 0.9003R_{\text{Lichess}}.\end{aligned}$$

Ove linije su postavljene preko ZKSKPROT3ZKSK ocena 650–1600 i ne bi trebalo da budu ekstrapolirano kao univerzalne konverzije na svim nivoima veština.

1 Cilj

Cilj je mapirati ocenu na ZKSKPROT0ZKSK na uporedivu skalu na ZKSKPROT2ZKSK za istu široku kategoriju kontrole vremena. Procenjujemo odvojeno mapiranje za svaku kategoriju koju koristi ZKSKPROT1ZKSK:

- 1+0 metak,
- 3+0 blic,

- 10+0 brzi,
- brze igre duže od 10 minuta.

Uklopljena linija se koristi samo kao sloj za konverziju skale. To ne podrazumeva to dva rejtinga imaju identičnu mehaniku ocenjivanja, grupe igrača ili ponašanje izbora igre.

2 Data Provenance

ZKSKPROT3ZKSK strana analize napravljena je od kompletnog javnog ZKSKPROT3ZKSK arhive igara sa standardnom ocenom od januara 2025. do marta 2026. Ovo je važno jer ZKSKPROT3ZKSK uzorak nije praktičan ili uski uzorak vidljivi profili: napravljen je od kompletnih mesečnih fajlova istorije igara izdao ZKSKPROT3ZKSK za istraživanje i javnu upotrebu.

ZKSKPROT3ZKSK objavljuje ove izvoze baze podataka na ZKSKPROT12ZKSK. Stranica baze podataka navodi da su izvozi objavljeni pod "ZKSKPROT11ZKSK licencom" i eksplicitno dozvoljava korisnicima da ih "Koriste za istraživanje" i "publication". ZKSKPROT13ZKSK Ovi otvoreni podaci držanje je centralno za dizajn referentne strane ZKSKPROT1ZKSK.

Na strani ZKSKPROT2ZKSK, proveravaju se samo kandidati sa istim korisničkim imenom mesečne arhive i kategorije igara potrebne za procenu mapiranja. Mi stavite poseban naglasak na podudaranje istorije utakmica u okviru istog kalendara meseca kad god je to moguće. Podudaranje kalendarskog meseca smanjuje rizik da a rejting igrača se značajno promenio između ZKSKPROT3ZKSK i ZKSKPROT2ZKSK merenja.

3 Konstrukcija utakmica kandidata

Glavni izazov je što nijedna platforma ne izlaže univerzalnu međuplatformu identifikator igrača. Stoga počinjemo sa kandidatima sa istim korisničkim imenom: ako a korisničko ime se pojavljuje i na ZKSKPROT0ZKSK i na ZKSKPROT2ZKSK, taj par se tretira kao utakmica kandidata. Ovo je namerno bučno. Neki nalozi sa istim imenom su ista osoba, dok su drugi nepovezani igrači koji slučajno koriste istu ručka.

Da bi se smanjila buka koja se može izbeći, trenutni cevovod se primenjuje specifično za tip igre filteri:

1. Korisnici ZKSKPROT3ZKSK su uzorkovani iz stratifikovanih skupova podataka o istoriji igara od strane ocenjivanje i tip kontrole vremena.
2. ZKSKPROT2ZKSK nalozi kandidata se proveravaju da li postoje odgovarajući kategorija kontrole vremena.
3. Parovi kandidata se zadržavaju samo kada obe strane imaju dovoljno relevantnih istorija igre.
4. Za metak, blitz i 10-minutni brzi, minimalna zadržana istorija je 5 utakmica po strani.
5. Za brzi zapis duži od 10 minuta, minimalna zadržana istorija je 10 igre po strani.
6. Regresijsko uklapanje prijavljeno ovde je ograničeno na ZKSKPROT3ZKSK ocene od 650 do 1600, što odgovara regionu rejtinga koji je najrelevantniji za glavni ZKSKPROT1ZKSK benchmark izveštaji.

Trenutni uzorak naglašava podudaranja istog korisničkog imena za isti kalendar i mesec. Kada je a nalog kandidata ima dovoljno igara u kategoriji ciljne kontrole vremena u istog meseca kao i izvorni podaci ZKSKPROT3ZKSK, taj odgovarajući mesečni rejting je preferirano za uklapanje. Stari podaci o podudaranju na više platformi se zadržavaju tamo gde su povećati podršku uzorku, ali metodološki cilj je specifičan za tip igre poređenje istog meseca.

4 Zašto srednja ili srednja vrednost mogu biti pristrasna

Za fiksnu kategoriju ZKSKPROT3ZKSK, skup kandidata za isto korisničko ime je mešavina:

$$f(y | x) = \pi f_{\text{true}}(y | x) + (1 - \pi) f_{\text{false}}(y),$$

gde je ZKSKPROT12ZKSK(kZKSKPROT13ZKSK) ocena ZKSKPROT3ZKSK, ZKSKPROT14ZKSK(iZKSKPROT je ocena ZKSKPROT2ZKSK, f_{true} is the distribution of Chess.com ratings for true podudaranja na više platformi, a ZKSKPROT12ZKSK(f_ZKSKPROT13ZKSKZKSKPROT14ZKSK) je generisana distribucija od strane nepovezanih ZKSKPROT2ZKSK naloga sa istim korisničkim imenom.

Komponenta lažnog podudaranja nije usredsređena na odgovarajući ZKSKPROT2ZKSK ekvivalent za kašiku ZKSKPROT3ZKSK. Ima tendenciju da liči na široku populaciju ZKSKPROT2ZKSK uzorkovano korisničkim imenima i filterima aktivnosti. Kao rezultat toga, uslovna sredina je povučeno ka globalnoj distribuciji aktivnih igrača:

$$E[y | x] = \pi E_{\text{true}}[y | x] + (1 - \pi) E_{\text{false}}[y].$$

Ako ZKSKPROT12ZKSK(ZKSKPROT13ZKSK < 1ZKSKPROT14ZKSK), ovo generalno nije jednako željenom podudarnom igraču očekivanje. Isti problem utiče na medijanu kad god je lažno podudaranje komponenta doprinosi dovoljnoj masi da pomeri 50. percentil.

Pravac pristrasnosti zavisi od kante. U niskim ZKSKPROT3ZKSK kantama, netačno podudaranja mogu povući srednju vrednost nagore ako su nasumični aktivni ZKSKPROT2ZKSK nalozi jači od pravih podudarnih korisnika. U visokim kantama, ista kontaminacija može povući srednju vrednost nadole. Ovaj efekat kompresije je upravo ono što želimo izbegavajte pri izgradnji konverzije razmera.

5 Modalna motivacija

The modal approach is based on a weaker assumption than the mean-based approach. We do not require every same-username pair to be correct. Instead, we assume that within a sufficiently narrow Lichess rating bucket, true matches form the most concentrated local cluster in Chess.com rating space. False matches may be numerous, but they should be more diffusely distributed because they are not tied to the Lichess rating bucket.

Za svaku kantu ZKSKPROT3ZKSK procenjujemo način ocene ZKSKPROT2ZKSK raspodela korišćenjem procene gustine jezgra. Zatim se tretira modalna tačka kao najverovatniji ZKSKPROT2ZKSK ekvivalent za tu ZKSKPROT3ZKSK kantu. Ovo čini postupak manje osetljiv na duge repove i manje osetljiv na nepodudaranja nalozi sa istim imenom uzorkovani iz šire populacije ZKSKPROT2ZKSK.

Ovo nije tvrdnja da je režim savršen. To je pragmatična robusnost izbor za bučan problem rešavanja entiteta. Metoda najbolje funkcioniše kada je istinita podudaranja ostaju množina blizu tačne ocene na više platformi i lažna šibice same po sebi nisu koncentrisane na pogrešnu vrednost.

6 Procedura ugradnje

Trenutna proizvodnja koristi sledeću proceduru posebno za svaku vrsta igre:

1. Zadržite parove kandidata sa istim korisničkim imenom koji zadovoljavaju broj igara pragovi.
2. Ograničite uklapanje na ZKSKPROT12ZKSK(650 ZKSKPROT13ZKSKR_ZKSKPROT14ZKSK ZKSKPROT15ZKSK1600ZKSKPROT16ZKSK).
3. Podeliti ZKSKPROT3ZKSK ocene u segmente od 100 tačaka.
4. Za svaku kantu sa najmanje 30 korisnika procenite ZKSKPROT2ZKSK modal rejting sa procenom gustine Gausovog jezgra koristeći ocenu propusnog opsega 75 bodova.
5. Predstavite svaku kantu po njenoj sredini i procenjenom režimu ZKSKPROT2ZKSK.
6. Postavite običnu liniju najmanjih kvadrata kroz modal na nivou korpe tačke:

$$\hat{R}_{\text{Chess.com}} = \alpha + \beta R_{\text{Lichess}}.$$

Samo modalne tačke na nivou segmenta ulaze u konačnu linearnu regresiju. Pojedinač parovi istog korisničkog imena utiču na krivu kroz svoj doprinos procena gustine unutar kante.

7 Current Estimates

Tabela ZKSKPROT12ZKSK prikazuje trenutne ugrađene konstante i nagibe. Ovo su vrednosti koje su trenutno povezane u ZKSKPROT1ZKSK nakon 29. maja 2026. ažurirati.

Game type	Total users	Baseline	Same-month	In range	α	β
1+0 bullet	2,999	1,474	1,525	1,843	-530.60	0.9859
3+0 blitz	3,101	1,306	1,795	2,075	-548.88	1.0837
10+0 rapid	3,084	1,509	1,575	2,523	-499.68	1.0778
> 10 rapid	750	445	305	575	-351.72	0.9003

Table 1: Trenutni modalni linearni odgovara ZKSKPROT3ZKSK ocenama 650–1600. “U opseg” broji korisnike zadržane unutar ugrađenog ZKSKPROT3ZKSK intervala. Osnovni redovi dolaze iz postojeće tabele sa tačnom kontrolom među platformama; dolaze redovi istog meseca iz mesečnog uzorka istog korisničkog imena iz marta 2026.

Brza procena duža od 10 minuta ima značajno manju podršku od ostale tri kategorije. Stoga je najverovatnije da će se promeniti kao dodatni prikupljeni su podudarni korisnici.

8 Interpretacija

Uklopljene jednačine se najbolje tumače kao lokalne konverzije za početnike rani napredni korisnici, a ne kao univerzalni identiteti ocenjivanja. Na primer, the 10-minutna brza procena je

$$\hat{R}_{\text{Chess.com}} = -499.68 + 1.0778 R_{\text{Lichess}}.$$

Cross-platform same-username rating matches and modal mapping fits

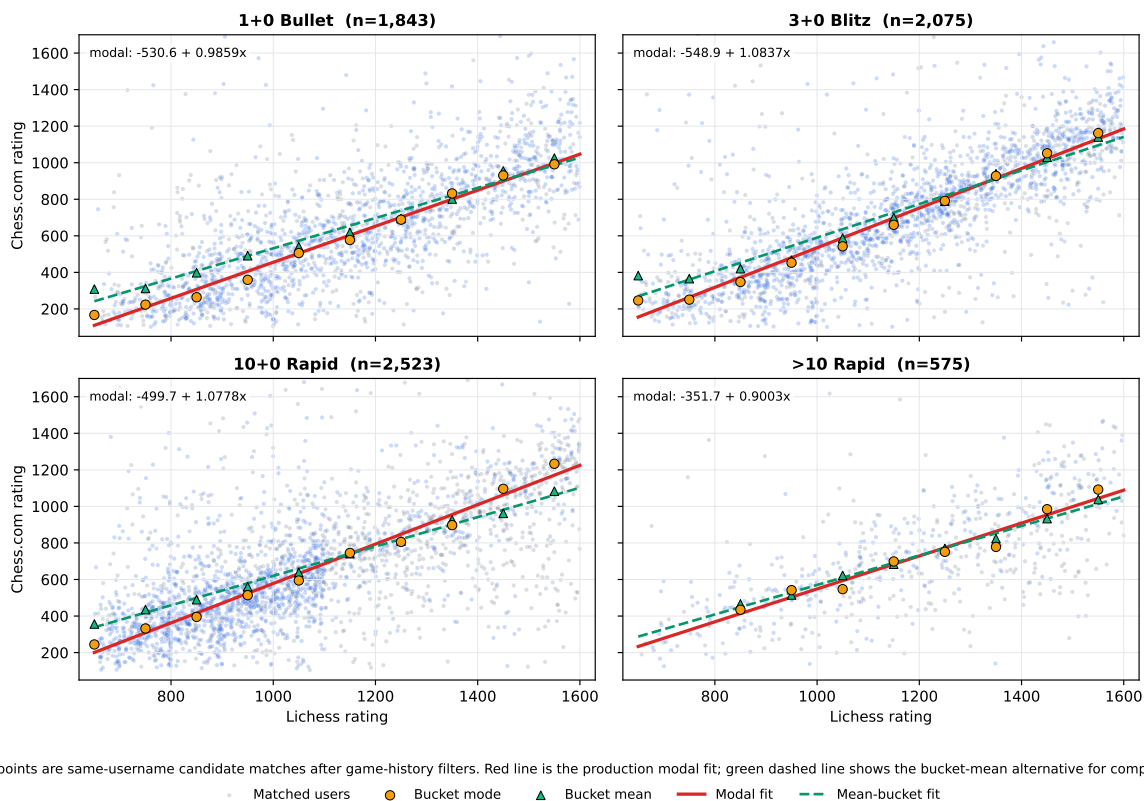


Figure 1: Podudaranje kandidata sa istim korisničkim imenom i uklopljene krive mapiranja. Sirovi bodovi su zadržani korisnici kandidati nakon filtera istorije igara. Narandžaste tačke su ZK-SKPROT2ZKSK ocene koje se koriste za proizvodno uklapanje. Crvena linija je linija modalne regresije. Pokazuju se zeleni trouglovi i isprekidana zelena linija alternativa srednje vrednosti za poređenje; ovi su uključeni radi ilustracije kako se rezimei zasnovani na srednjim vrednostima mogu izvući neusklađenim nalogima istog imena i zapažanja repa.

Sa ocenom ZKSKPROT3ZKSK od 1200, ovo daje ZKSKPROT2ZKSK ekvivalent od oko 793. Sa ocenom ZKSKPROT3ZKSK od 1500, daje oko 1117.

Procene ne treba ekstrapolirati daleko izvan prilagođenog intervala. At veoma visoke ocene, sastav grupe igrača, ponašanje naloga i platforma mehanika ocenjivanja može da se razlikuje na načine koji nisu obuhvaćeni jednim linearnim mapiranje prilagođeno nižem i srednjem rangu.

9 Ograničenja i plan ažuriranja

Sadašnja procedura je namerno konzervativna, ali ima nekoliko ograničenja ostati:

- Podudaranje istog korisničkog imena je bučno i ne dokazuje identitet.
- Deo pravog podudaranja može da varira u zavisnosti od ranga i tipa igre.
- ZKSKPROT2ZKSK i ZKSKPROT3ZKSK ocene se ne mere uvek na isti način dan za sve parove kandidata, iako uzorak istog meseca ovo smanjuje pitanje.
- Brzi uzorak duži od 10 minuta je trenutno manji od metak, blitz i 10-minutni brzi uzorci.
- Linearni model je praktična aproksimacija preko 650–1600, a ne a strukturni model bilo kog sistema ocenjivanja.

Sledeće planirano poboljšanje je nastavak uzorkovanja dužih brzih korisnika, podleže konzervativnim uslovima korišćenja i platforme ZKSKPROT10ZKSK. Jednom podudarni uzorak je veći i izbalansirani u svim segmentima rejtinga, modalna uklapanja bi trebalo da budu ponovo procenjena i proizvodne konstante ponovo ažurirane.

10 Održavanje i ažuriranja

Ovaj dokument odražava najnovije ZKSKPROT1ZKSK mapiranje ocena na više platformi ažurirano od 29. maja 2026. Procene će biti revidirane kao dodatne prikupljaju se korisnici koji se podudaraju sa istim kalendarskim mesecom, posebno za brze igre duže od 10 minuta.

Održana javna verzija ovog izveštaja dostupna je na:

<https://www.elopluschess.com/static/docs/cross-platform-elo-mapping-modal-method.pdf>