

Statističko zaključivanje o egzogenom i endogenom širenju informacija u društvenim mrežama

Matija Piškorec



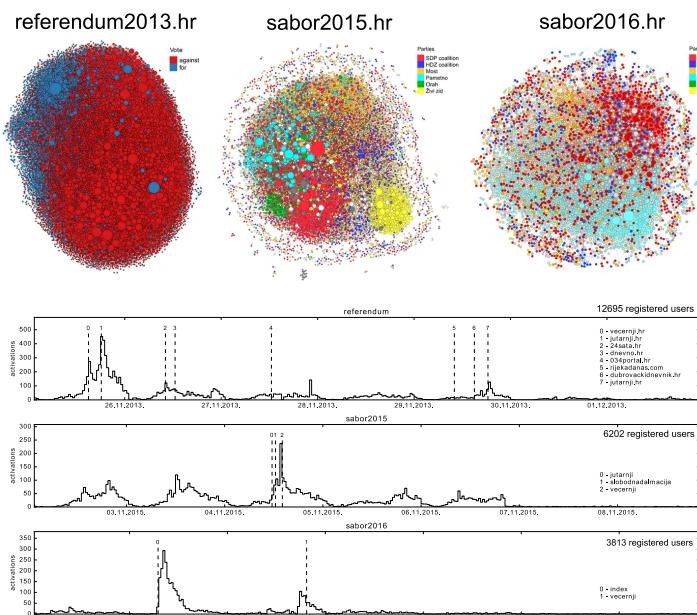
mentor: izv. prof. dr. sc. Mile Šikić¹, dr. sc. Tomislav Šmuc²

¹ Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva

² Institut Ruđer Bošković

1. Uvod

Ljudska interakcija se sve više odvija preko online informacijskih sustava kao što su blogovi, društvene mreže i servisi za dijeljenje sadržaja. Ovo nam daje jedinstvenu priliku da mjerimo i modeliramo društveni utjecaj na razinu koja je bila nezamisljiva prije par desetljeća. Utjecaj između korisnika online informacijskih sustava se može karakterizirati i modelirati preko interakcija koje međusobno ostvaruju. Cilj ovog doktorskog istraživanja je razvoj modela utjecaja u društvenim mrežama koji uključuje utjecaj između korisnika (endogeni utjecaj) ali i utjecaj van društvene mreže (egzogeni utjecaj). Predlažemo efikasnu statističku metodologiju koja može procijeniti iznos ova dva utjecaja iz podataka o aktivnosti korisnika na društvenoj mreži [3]. Evaluaciju modela i metodologije provodimo na podacima o aktivnosti korisnika na sruštvenoj mreži Facebook.



2. Metodologija

Estimaciju endogenog i egzogenog utjecaja u mreži provodimo pomoću optimizacije log-izglednosti sljedeće funkcije izglednosti:

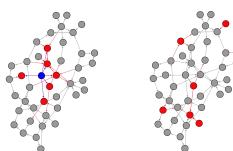
$$\mathcal{L}(D, t; p_{peer}(t), p_{ext}(t)) = \prod_{i \in \text{active}} (1 - (1 - p_{peer}^{(i)}(t))(1 - p_{ext}^{(i)}(t))) \prod_{i \in \text{nonactive}} (1 - p_{peer}^{(i)}(t))(1 - p_{ext}^{(i)}(t))$$

Prepostavka je da vjerojatnost aktivacije zbog endogenog utjecaja ovisi o već aktiviranim korisnicima u mreži, dok vjerojatnost aktivacije zbog egzogenog utjecaja je neovisna od toga i jednaka je za sve korisnike (iako se njen iznos može mijenjati kroz vrijeme). Za endogeni utjecaj odabiremo nekoliko modela koji se tipično koriste u literaturi [1]: (i) Susceptible-infected (SI) model, (ii) Exponential decay (EXP) model, i (iii) Logistic threshold (LOG) model:

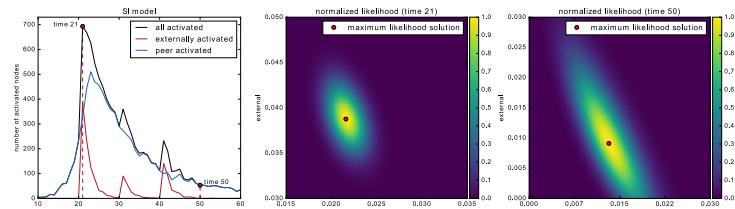
$$\text{SI: } p_{peer}^{(i)}(t) = 1 - \prod_{j \in \text{active}} (1 - p_0 e^{-\lambda(t-t_j)})$$

$$\text{EXP: } p_{peer}^{(i)}(t) = \frac{1}{1 + e^{-k(a_i(t) - a_0)}}$$

$$\text{LOG: } p_{peer}^{(i)}(t) = 1 - \prod_{j \in \text{active}} (1 - p_{SI}) = 1 - (1 - p_{SI})^{a_i(t)}$$

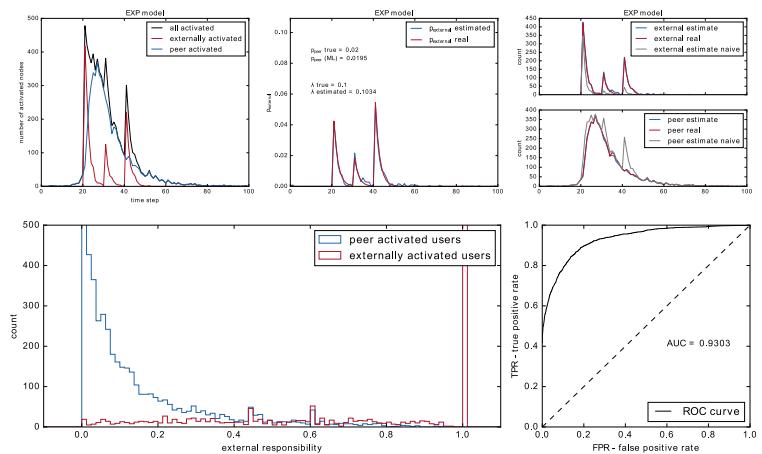


Ova izglednost se može optimizirati u svakom vremenskom prozoru kako bi se dobili parametri maksimalne izglednosti za endogeni i egzogeni utjecaj. Slika dolje pokazuje oblik izglednosti i log-izglednosti za simuliranu aktivacijsku kaskadu koja slijedi SI model endogenog utjecaja s dodanim egzogenim utjecajem koji se mijenja kroz vrijeme.

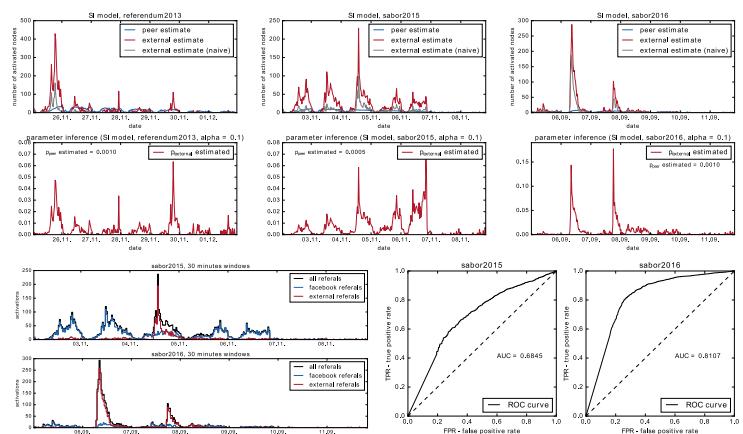


3. Rezultati

Našu metodologiju prvo evaluiramo na simuliranim informacijskim kaskadama na društvenoj mreži iz referendum2013.hr skupa podataka. Endogeni utjecaj modeliramo pomoću EXP modelu, dok egzogeni utjecaj dizajniramo kao eksponencijalno opadajući šiljak što otprilike odgovara egzogenom utjecaju iz vanjskih medija koje smo dobili eksploratornom analizom sabor2015.hr i sabor2016.hr skupova podataka. Za estimaciju jedinstvenog endogenog utjecaja kroz cijeli vremenski raspon koristimo alternirajući metodu sličnu onoj koja se već pojavljuje u literaturi [2]:



Našu metodologiju smo evaluirali i na stvarnim aktivacijskim kaskadama prikupljenima na društvenoj mreži Facebook, pretpostavljajući SI model za endogeni utjecaj i koristeći vremenske prozore od 30 minuta za estimaciju izglednosti. Egzogeni utjecaj estimiramo za svaki vremenski prozor zasebno.



4. Zaključak

Našom metodologijom moguće je, koristeći samo podatke o vremenu aktivacije korisnika društvene mreže i njihove mreže prijateljstava, karakterizirati aktivaciju svakog pojedinačnog korisnika kao i estimirati ukupni iznos endogenog i egzogenog utjecaja na aktivaciju svih korisnika.

5. Literatura

- [1] A. Anagnostopoulos, G. Brova, and E. Terzi. "Peer and authority pressure in information-propagation models", Proceedings of the ECML/PKDD 2011
- [2] S. Myers, C. Zhu, and J. Leskovec. "Information diffusion and external influence in networks", In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2012)
- [3] M. Piškorec, N. Antulov-Fantulin, I. Miholić, T. Šmuc, M. Šikić. "Modeling peer and external influence in online social networks: Case of 2013 referendum in Croatia", Complex Networks and Their Applications VI. Studies in Computational Intelligence, vol. 689, Springer (2017)

6. Zahvale

