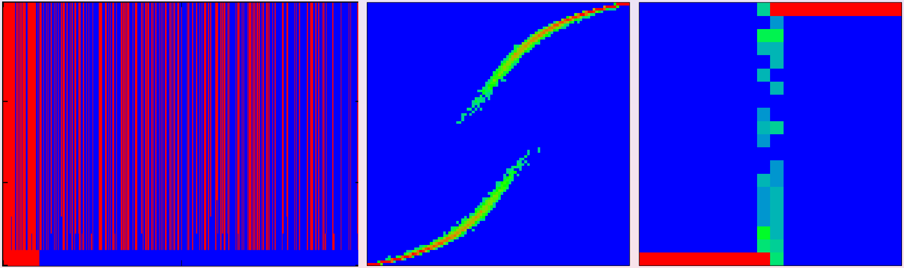
PageRank model of opinion formation

arXiv:1204.3806 (to appear in physica A)

Vivek Kandiah in collaboration with Dima Shepelyansky (CNRS, Toulouse)
www.quantware.ups-tlse.fr

Supported by Région Midi-Pyrénées, EC FET open project NADINE



Laboratoire de Physique Théorique, IRSAMC, UMR 5152 du CNRS
Université Paul Sabatier, Toulouse

Social Networks Today

- World Wide Web : Rapidly growing network, $\sim 10^{10}$ sites

  (http://www.worldwidewebsize.com/)
- Socio-physics : understanding social phenomena, mainly opinion dynamics
- Social Networks (Facebook, Twitter, VKONTAKTE, LiveJournal,...) : Sharing social and political views, hundreds of millions users, features of real networks, study of mass opinion formation
- Cambridge University website Network $N = 212710$, $N_l = 2015265$
- Oxford University website Network $N = 200823$, $N_l = 1831542$ (2006)

  Academic Weblink Database Project, http.cybermatrices.wlv.ac.uk/database/.
- LiveJournal $N = 3577166$, $N_l = 44913072$
- Twitter $N = 41652230$, $N_l = 1468365182$

Benczur (2008)
Vigna : http://vigna.dsi.unimi.it/

Galam (2008)
Castellano,Fortunato,Loreto (2009)
Kwak,Lee,Park,Moon (2010)

Elector networks

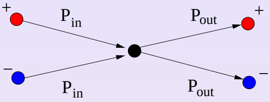
- Many studies of opinion formation on regular lattices, voter model, Sznajd model, etc.
- Real social networks show small-world and scale-free properties
- PageRank is an efficient ranking technique and provides a natural order of importance in a network
- PageRank top nodes represent the elite among the social network

Idea : One's opinion is influenced by the closest members (friends) among the society and influential friend's opinion count more than less important friend's opinion in our environment.

Implementation : Two possible opinions coded by Ising spin variables σ_i taking values 1 or -1. We choose an initial distribution of opinions on the network and observe how the fraction of nodes having the same opinion evolves during time according to a certain rule.

Holley and Liggett (1975)
Sznajd-Weron (2000,2002,2004,2005)
Krapivsky,Redner,Ben-Naim (2010)

PageRank Opinion Formation Model (PROF)



$$\Sigma_i = a\sum_j P^+_{j,in} + b\sum_j P^+_{j,out} - a\sum_j P^-_{j,in} - b\sum_j P^-_{j,out} \;\; , \;\; a+b=1$$

P_j : PageRank of node j

- defined for one iteration step
- σ_i takes the value 1 or -1 respectively for $\Sigma_i > 0$ or $\Sigma_i < 0$.
- The parameters a and b allow to tune the importance of incoming and outgoing links.

Large b $\rightarrow$ an elector takes the opinion of people he is looking at $\rightarrow$ "conformist" society

Large a $\rightarrow$ an elector takes mainly the opinion of people pointing to him $\rightarrow$ "tenacious" society

Time evolution of opinion fractions

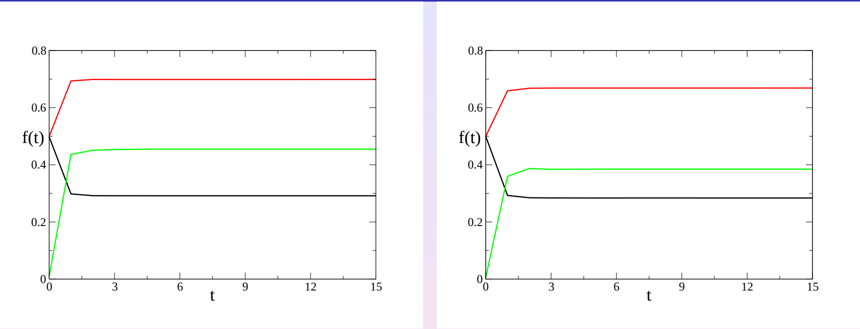
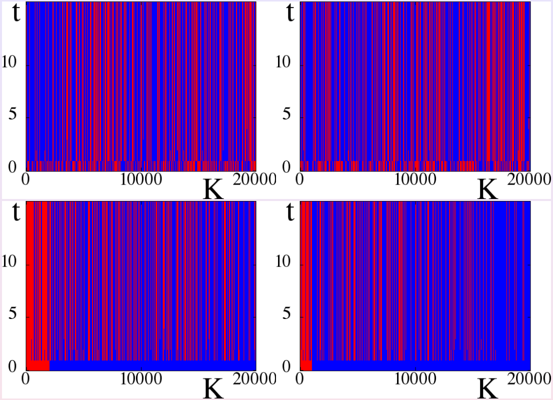


Figure: Time evolution of opinion (in term of fraction or red nodes) as function of number of iterations t for Cambridge (left, $N_{top} = 2000$) and Oxford (right, $N_{top} = 1000$). The green curves show the evolution of opinion when N_{top} nodes are red. Here $a = b = 0.5$.

- sign of bistability
- convergence to a fixed state, in a time O(1) as on regular lattices
- Top rank nodes can impose their opinion to a significant fraction of nodes
- Cheirank ineffective

In PROF model, the elite can influence significantly the whole society if they have fixed opinions between themselves

Time evolution of opinion fractions



Corresponding time evolution of colour opinion with same parameters for Cambridge (left) and Oxford (right). Down panels show the evolution of opinion when N_{top} nodes are red. The final red nodes are homogeneously distributed in K (index of decreasingly ordered PageRank probabilities).

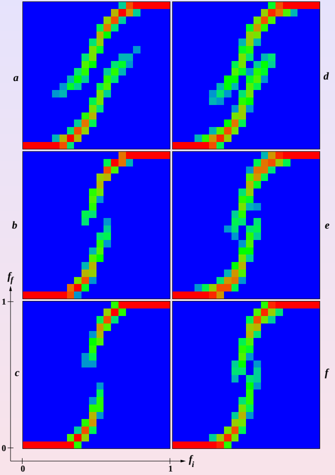


Figure: Density plots of probability to find a final fraction f_f depending on initial fraction f_i of red nodes for Cambridge (left) and Oxford (right). $N_r = 10^4$ random realizations (up to convergence time t=20 iterations) were used on a 20x20 cells grid. From top to bottom $a = 0.1$,$a = 0.5$ and $a = 0.9$.

- small fraction of red opinion suppressed/larger fraction dominates
- range of bistability phase, wider for low a

A tenacious society has a relatively small range of bistability phase unlike the conformist society where the opinion is strongly influenced by elite.
random initial distribution of opinion $\rightarrow$ divided elite $\rightarrow$ divided followers $\rightarrow$ large bistable region

Influence of the society elite

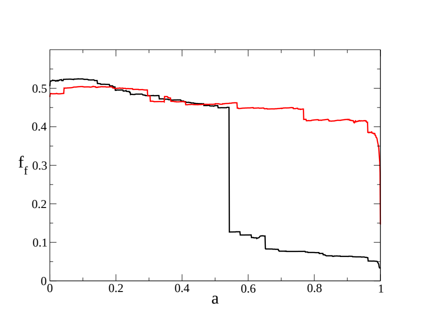
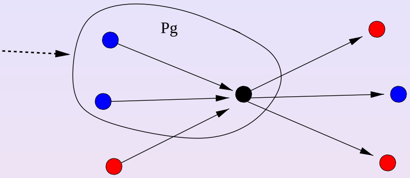


Figure: Dependance of f_f on a for $N_{top} = 2000$ red nodes, for Cambridge (black) and Oxford (red).

- a=0, the society follows in majority the opinion of the elite
- a=1, the final fraction drops $\rightarrow$ "tenacious" society

For small values of a, the main influence are the outgoing links and since the PageRank probability is proportional to the number of ingoing links, the nodes having a lot of incoming links are the elite. In the limit of small a, the society members form their opinion listening to an elite opinion. If the elite has the same opinion among themselves, it can easily impose it to a large fraction of the society.

PROF-Sznajd model



A group point of view describing the famous principle :

"United we stand, divided we fall"

- pick a random node $\rightarrow$ polarization of N_g -1 highest PageRank nodes pointing to it ?
- if they have the same polarization $\rightarrow$ group with effective PageRank $P_g = \sum_{j=1}^{N_g} P_j$
- consider all nodes pointing to any member of the group
- check all those n nodes, if $P_n < P_g$: the node joins the group by taking the same polarization and P_g is increased by P_n. (preventing small groups to influence high rank members).

Features of society described by PROF-Sznajd model

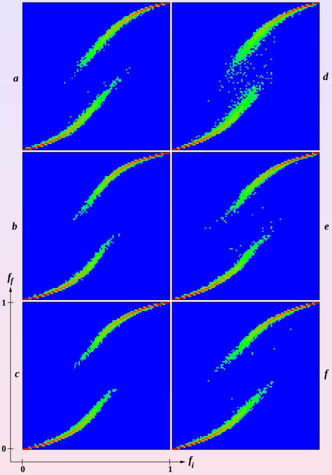


Figure: Density plot of probability constructed using $N_r = 10^4$ random realizations following the evolution up to the convergence time $\tau = 10^7$ iterations for Cambridge (left) and Oxford (right). Here from top to bottom $N_g = 3$, $N_g = 8$ and $N_g = 13$.

- bistability phase
- smaller fluctuations at larger N_g
- finite f_f at small $f_i \to$ resistance of small groups against totalitarian opinion

Opinion formation in LiveJournal network

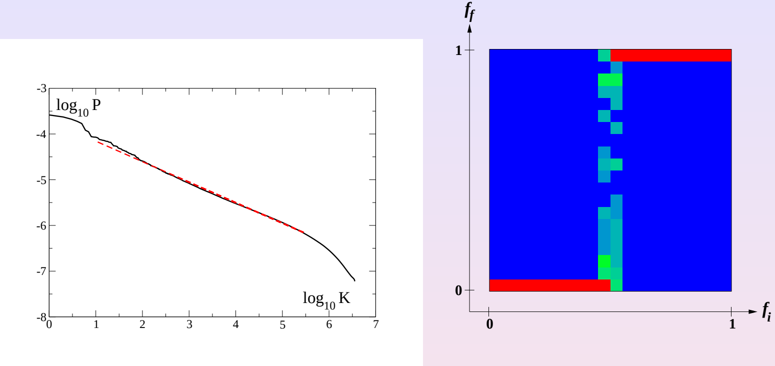
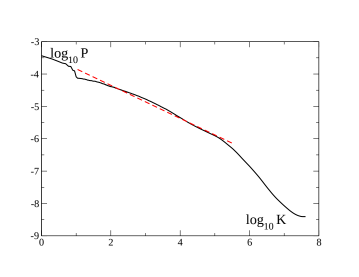


Figure: Left panel : PageRank probability decay with index K, the red curve is the fitted algebraic dependance. Right panel : Density plot of probability for LiveJournal at $a = 0.5$.

- $N = 3577166$ nodes, $N_l = 44913072$ mainly directed links
- slower PageRank decay, $P(K) \propto 1/K^{\beta}$ with $\beta = 0.448 \pm 0.000046$
- similar convergence time scale $\sim$ O(1)
- bistability disappeared

Benczur (2008)

Opinion formation in Twitter network



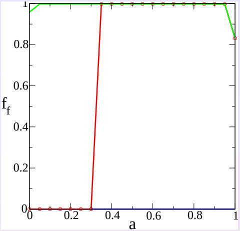


Figure: Left panel : PageRank probability decay with index K, the red curve is the fitted algebraic dependance. Right panel : Dependance of f_f on tenacious parameter a in PROF model for initial N_{top} red nodes. Here $N_{top} = 1200$ (blue), $N_{top} = 1250$ (red) and $N_{top} = 1300$ (green).

- $N = 41652230$ nodes and $N_l = 1468365182$ links. Decay exponent $\beta = 0.51$ for $1 \leq log_{10}K \leq 5.5$ and $\beta = 1.23$ for $5.5 \leq log_{10}K \leq 7$
- small fraction of elite ($N_{top}/N \approx 3 \cdot 10^{-5}$) can impose its opinion practically to the whole society for all values of a
- very connected network, large average number of links per node $\rightarrow$ sharp transition

Kwak,Lee,Park,Moon (2010), Vigna

References

1. J.R. Zaller, *The Nature and origins of mass opinion*, Cambridge University Press, Cambridge UK, 1999.
2. S. Galam, J. Math. Psychology **30** (1986) 426.
3. T.M. Liggett, *Stochastic interacting systems: contact, voter and exclusion processes*, (Springer, Berlin) 1999.
4. S. Galam, Europhys. Lett. **70** (2005) 705.
5. S. Galam, Int. J. Mod. Phys. C **19** (2008) 409.
6. C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81** (2009) 591.
7. P. L. Krapivsky, S. Redner and E. Ben-Naim, *A Kinetic view of statistical physics*, Cambridge University Press, Cambridge UK, 2010.
8. Academic Web Link Database Project `http://cybermetrics.wlv.ac.uk/database/|`
9. M. Kurucz, A.A. Benczur, A. Pereszlenyi, *Large-Scale Principal Component Analysis on LiveJournal Friends Network*, Proc. *Workshop on Social Network Mining and Analysis* held in conjunction with *13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008)*, Las Vegas NV, August 24-27 (2008); `http://dms.sztaki.hu/en/letoltes/livejournal-data|`
10. H. Kwak, C.Lee, H. Park and S. Moon, *What is Twitter, a social network or a news media?*, Proc. 19th Int. Conf. WWW2010, p.591, ACM, New York, N.Y. 2010; the web data are downloaded from the web site maintained by S.Vigna http://law.dsi.unimi.it/webdata/twitter-2010/
11. S.Brin and L.Page, *Computer Networks and ISDN Systems* **30** 107 (1998).
12. A.M. Langville and C.D. Meyer C D 2006 *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, Princeton, 2006.
13. S. Redner, Phys. Today **58(6)** (2005) 49.
14. F. Radicchi, S. Fortunato, B. Markines, and A. Vespignani, Phys. Rev. E **80** (2009) 056103.
15. T. Preis, D. Reith and H.E. Stanley, Phil. Trans. R. Soc. A **368** (2010) 5707.
16. T. Preis, H.S. Moat, H.E. Stanley and S.R. Bishop, Sci. Reports **2** (2012) 350.
17. K. Sznajd-Weron and J. Sznajd, Int. J. Mod. Phys. C **11** (2000) 1157.
18. K.M. Frahm, B. Georgeot and D.L. Shepelyansky, J. Phys, A: Math. Theor. **44** (2011) 465101.
19. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H.Teller, and E. Teller, J. Chem. Phys. **21** (1953) 1087.
20. V. Sood and S. Redner, Phys. Rev. Lett. **94** (2005) 178701.
21. N.S. Ananikian and S.K. Dallakian, Physica D **107** (1997) 75.