

Mathematical Models of Information and Communication Systems

Vladimir V. Mazalov

Institute of Applied Mathematical Research, Karelian Research
Center, Russian Academy of Sciences, Russia

February 13, 2017

<i>New Trends</i>	<i>Mathematical Theory</i>
<i>Internet and WWW</i> <i>Mobile communication</i> <i>Social networks</i> <i>Transportation networks</i> <i>e – economics e – business</i> <i>e – marketing e – library</i> <i>Grid and Super computing</i> <i>Cloud computing</i>	<i>Graph theory, Webometrics</i> <i>Probability theory and Statistics</i> <i>Game, Information, Reputation theory</i> <i>Logistic, Optimal routing</i> <i>Optimisation theory</i> <i>Multiagent systems, Data mining</i> <i>Algorithms, Complexity theory</i> <i>Resource allocation, Queueing theory</i>

Here is the weighted graph extracted from the popular Russian social network VKontakte. The graph corresponds to the online community devoted to game theory. This community consists of 483 participants. As a weight of a link we take the number of common friends between the participants.

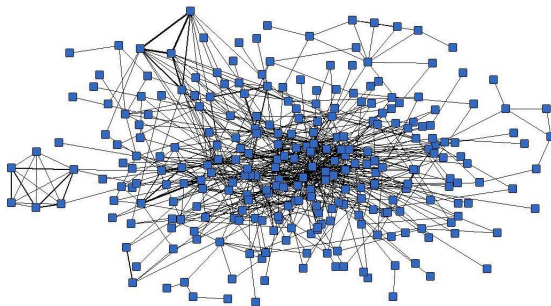
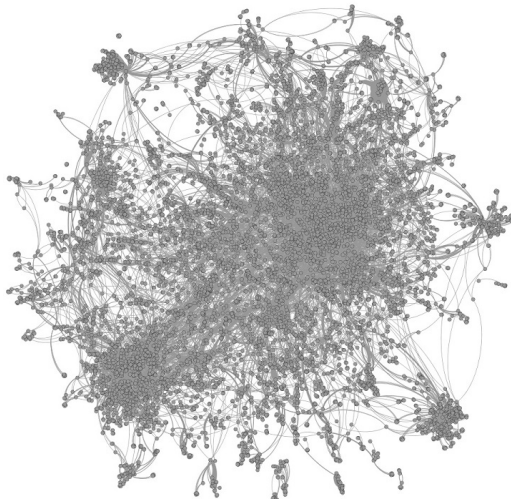
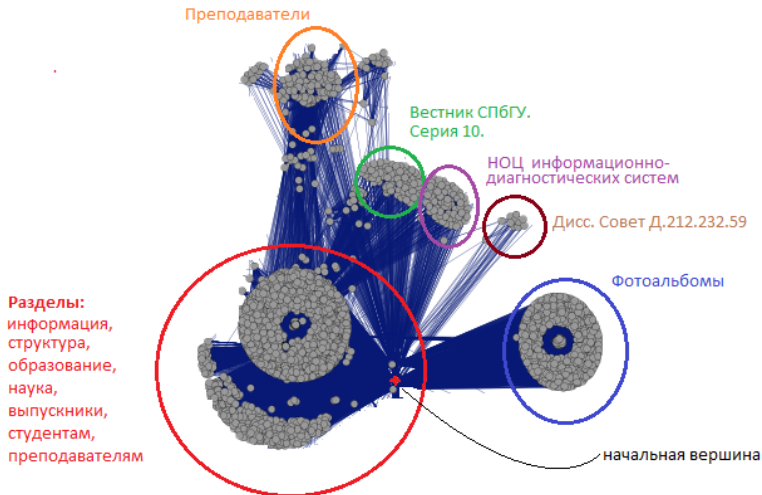


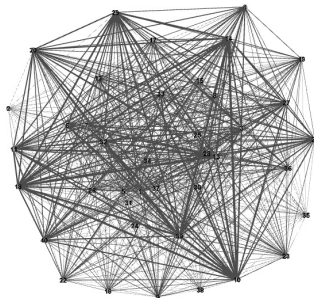
Figure: Principal component of the community Game Theory in the social network VKontakte (number of nodes: 275, number of edges: 805 and mean path's length: 3.36).

On fig. 2 it is presented the subgraph from the Russian mathematical portal Math-Net.ru. The general amount of the authors on the mathematical portal Math-Net.ru now is equal to 78839.





Graph of joint specialities at Russian Universities



- Ranking of nodes
- Ranking of edges
- Clustering and Community detection
- Closeness in network
- Comparison of networks
- Prediction of links
- Flows

Betweenness centrality

Consider a weighted graph $G = (V, E)$, where

- V is the set of nodes,
- E is the set of edges.

Betweenness centrality of node $v \in V$ [Freeman, 1977]:

$$c_B(v) = \frac{1}{n(n-1)} \sum_{s,t \in V} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}, \quad (1)$$

where $\sigma_{s,t}$ is the total number of geodesics (shortest paths) between nodes $s \in V$ and $t \in V$; $\sigma_{s,t}(v)$ is the number of geodesics between s and t that v lies on.

The complexity of the fastest algorithm [Brandes, 2001]:

- on weighted graphs is $O(n^3)$ or $O(n^2 \log n + nm)$ on a sparse weighted graphs,
- on unweighted graphs is $O(mn)$,

where $n = |V|$ is the number of nodes, $m = |E|$ is the number of edges.

Example

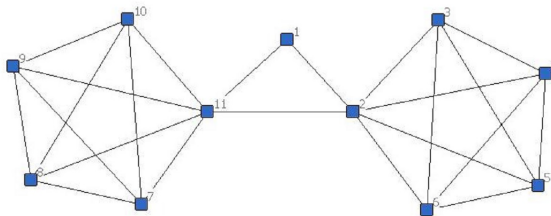


Figure: Network of 11 nodes, $c_B(1) = 0$

PageRank method was developed by [Brin, Page, 1998].

PageRank procedure works by counting the number of links to a web page to determine a random walk with matrix of transition probabilities P .

Let node v_i has $k > 0$ incoming links. Then $p_{ij} = 1/k$.

Markov chain is determined by

$$\tilde{P} = \alpha P + (1 - \alpha)(1/n)I,$$

where $\alpha \in (0, 1)$, I - is matrix of 1. In Google α is equal to 0.85. Matrix \tilde{P} is stochastic. Markov ergodic theorem \Rightarrow it exists vector π such that $\pi \tilde{P} = \pi, \pi \mathbf{1} = 1$. $\pi = \mathbf{PageRank}$.

For weighted graph with weight matrix W the transition matrix is $P = D^{-1}W$ where D is degree matrix. Transition matrix is

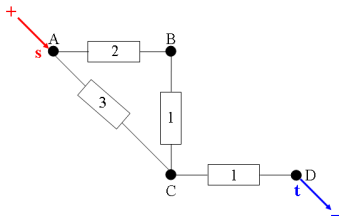
$$\tilde{P} = \alpha D^{-1}W + (1 - \alpha)(1/n)I.$$

Current flow centrality based on Kirchhoff's law

Consider a weighted undirected graph $G = (V, E, W)$, where

- V is the set of nodes
- E is the set of edges
- W is the matrix of weights:

$$W(G) = \begin{pmatrix} 0 & w_{1,2} & \dots & w_{1,n} \\ w_{2,1} & 0 & \dots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \dots & 0 \end{pmatrix}$$



We introduce the diagonal degree matrix:

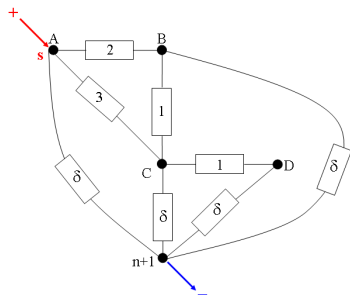
$$D(G) = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix},$$

where $d_i = \sum_{j=1}^n w_{i,j}$ is the sum of weights of the edges which are adjacent to node i in graph G . The Laplacian matrix $L(G)$ for weighted graph G is defined as follows:

$$L(G) = D(G) - W(G) = \begin{pmatrix} d_1 & -w_{1,2} & \dots & -w_{1,n} \\ -w_{2,1} & d_2 & \dots & -w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -w_{n,1} & -w_{n,2} & \dots & d_n \end{pmatrix}. \quad (2)$$

Current flow centrality based on Kirchhoff's law

Let the graph G' be converted from the graph G by extension with an additional node (grounded) $n + 1$ connected with all nodes of the graph G with the links of constant conductance β .



Thus, we obtain the Laplacian matrix for the modified graph G' as:

$$L(G') = D(G') - W(G') = \begin{pmatrix} d_1 + \beta & -w_{1,2} & \dots & -w_{1,n} & -\beta \\ -w_{2,1} & d_2 + \beta & \dots & -w_{2,n} & -\beta \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -w_{n,1} & -w_{n,2} & \dots & d_n + \beta & -\beta \\ -\beta & -\beta & \dots & -\beta & \beta n \end{pmatrix}. \quad (3)$$

Suppose that a unit of current enters into the node $s \in V$ and the node $n + 1$ is grounded. Let φ_i^s be the electric potential at node i when an electric charge is located at node s . The vector of all potentials $\varphi^s(G') = [\varphi_1^s, \dots, \varphi_n^s, \varphi_{n+1}^s]^T$ for the nodes of graph G' is determined by the **Kirchhoff's current law**:

$$L(G')\varphi^s(G') = b'_s, \quad (4)$$

where b'_s is the vector of $n + 1$ components with the values:

$$b'_s(i) = \begin{cases} 1 & i = s, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Current flow centrality based on Kirchhoff's law

The current let-through the link $e = (i, j)$ according to Ohm's law is $x_e^s = |\varphi_i^s - \varphi_j^s| \cdot w_{i,j}$. Define the β CF-centrality of edge e as

$$CF_\beta(e) = \frac{1}{n} \sum_{s \in V} x_e^s. \quad (6)$$

Given that the electric charge is concentrated at node s , the mean value of the current flowing through node i is

$$x^s(i) = \frac{1}{2}(b_s(i) + \sum_{e: i \in e} x_e^s). \quad (7)$$

And finally, define the beta current flow centrality (β CF-centrality) of node i in the form

$$CF_\beta(i) = \frac{1}{n} \sum_{s \in V} x^s(i). \quad (8)$$

Weighted network of six nodes.

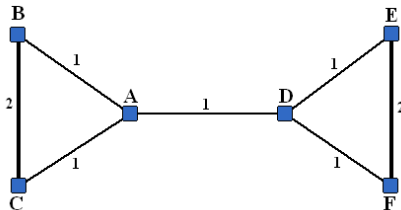


Figure: Weighted network of six nodes.

Classical betweenness centrality evaluates only the nodes A and D and gives 0 to other four nodes, even though they are obviously also important.

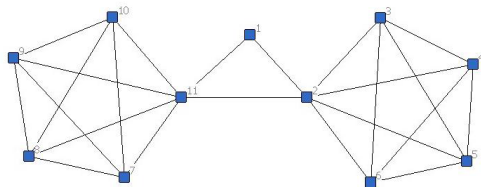
Table. Measures of centrality for weighted graph with six nodes.

Nodes	A	B	C	D	E	F
Original betweenness centrality	6	0	0	6	0	0
PageRank centrality $\alpha = 0.85$	1/6	1/6	1/6	1/6	1/6	1/6
Current flow betweenness centrality $\beta = 1$	0.27	0.19	0.19	0.27	0.19	0.19

The PageRank method ranks all nodes with equal values and thus it is indiscriminatory in this particular case.

The current flow betweenness centrality gives rather high values to nodes A and D.

Unweighted network of eleven nodes.



Nodes	Centrality $\beta = 0.5$	Edges	Centrality $\beta = 0.5$
2, 11	0.291	(2,11)	0.137
1	0.147	(1,2), (1,11)	0.101
other	0.127	other	0.0647

The centrality of nodes 2 and 11 is twice as great as that of node 1. At the same time, the centrality of node 1 and adjacent edges exceeds the centrality of the other nodes and edges in the network.

The results of computer experiments with network of mathematical publications Math-Net.ru.

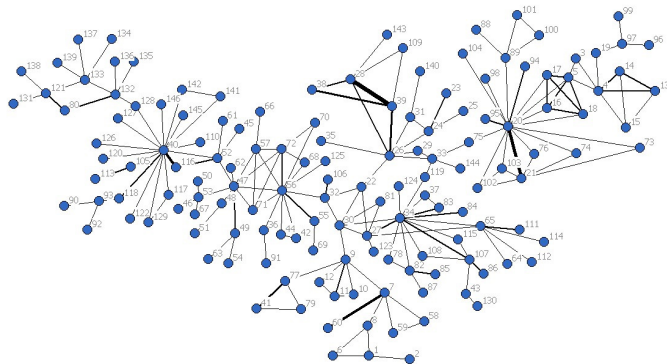


Figure: Graph from mathematical portal Math-Net.ru.

Graph contains 7606 authors and 10747 articles.

Results of ranking for the mathematical portal Math-Net.ru.

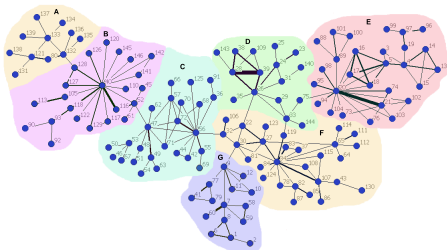
Node	Centrality (CF_{β})	Node	PageRank $\alpha = 0.85$	Node	CF-betweenness centrality
40	0.15740	40	0.04438	56	0.54237
34	0.14981	34	0.03285	32	0.53027
20	0.13690	20	0.03210	47	0.48222
47	0.12566	56	0.02774	22	0.41668
56	0.12518	47	0.02088	33	0.41361
26	0.10880	39	0.01874	34	0.39517
30	0.09098	28	0.01824	30	0.39426
9	0.08149	21	0.01695	52	0.37421
33	0.08024	65	0.01632	40	0.36946
32	0.07959	26	0.01552	26	0.35259

Illustrative examples








Community structure of the subgraph associated with Math-Net.ru.

By removing edges have a high βCF centrality, the groups are separated from each other, and the underlying community structure of the network is revealed.

(32, 56), (9, 30), (47, 52), (20, 75), (22, 26), (34, 119), (128, 132), (9, 11), (4, 5), ...



The graph splits into 7 communities corresponding to different fields of mathematics, namely, coding, discrete mathematics, mathematical physics, functional analysis, algebra and topology, optimal control, and probability theory.

-  Aumann, R., Myerson, R.: Endogenous formation of links between players and coalitions: an application of the Shapley value, in: *The Shapley value*, Cambridge University Press, 1988, 175-191.
-  Avrachenkov, K.E., Mazalov, V.V., Tsynguev, B.T.: Beta Current Flow Centrality for Weighted Networks, In Proceedings of CSoNET 2015, LNCS v.9197, 2015, pp.216-227
-  Brandes, U., Fleischer, D.: Centrality measures based on current flow, in: *Proceedings of the 22nd annual conference on Theoretical Aspects of Computer Science*, 2005, 533-544.
-  Freeman, L.C.: A set of measures of centrality based on betweenness, *Sociometry*, **40**, 1977, 35-41.
-  Jackson, M.O., Wolinsky, J.: A strategic model of social and economic networks, *J. Econ. Theory*, **71**(1), 1996, 44-74.
-  Mazalov, V.V., Trukhina, L.I.: Generating functions and the Myerson vector in communication networks, *Discrete Mathematics and Applications* **24**(5), 2014, 295-303.
-  Myerson, R.B.: Graphs and cooperation in games, *Math. Oper. Res.*, **2**, 1977, 225-229.