VK Analysis

Network Analysis

Anastasia Ignatyeva, AISA

NATIONAL RESEARCH UNIVERSITY

«HIGHER SCHOOL OF ECONOMICS»

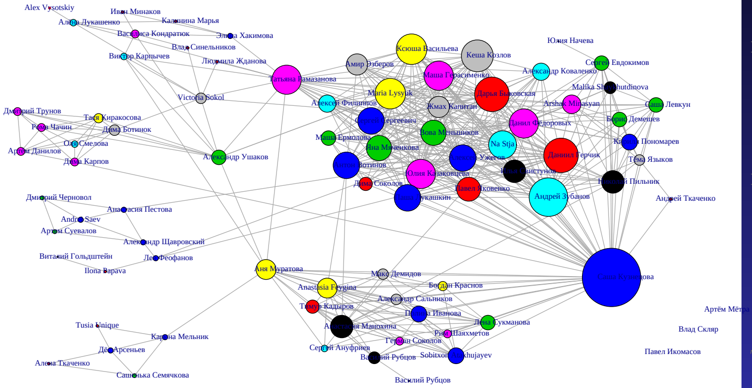
FACULTY OF COMPUTER SCIENCE

April, 2016

Contents

Graph of friends

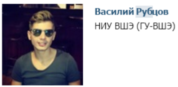
Graph : 85 nodes, 454 edges

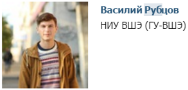


Who is my freind?

Graph : 85 nodes, 454 edges

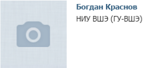
which includes:

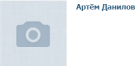
- two Vasiliy Rubtzov 

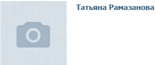
= 

- three friends with last name — «Scolov(a)» (not relatives)
- one friends with deleted account 

- three cage friends have no accounts' photo







Who is my friend?

Birthday:

- only day & month — 29
- no information — 18
- most popular date — the 3 of August (3 persons)

Relationship Status:

- not married — 13
- has boyfriend/girlfriend — 2
- fell in love — 7
- engaged — 1
- married — 1
- no information — 61

Faculty name:

- ecomonic science — 29
- computer science — 10
- physical — 2
- ...
- no information — 8

Sex:

- male — 54/85
- female — 31/85

Number of posts:

- max — 8303
- 6 individuals have no posts
- average value ≈ 1148

Days of inactivity:
(25.09.2016 - date of last post)

- max — 1302 (Nikolay Pilnik)
- average number ≈ 132
- (Alexander Salnikov, Danil Fedorovykh, Victoria Socol) — 0

University:

- NRU HSE — 46
- MSU — 7
- MIPT — 2
- ...
- no information — 8

Who is my friend?

Node Degree:

- max degree — 44 (Sasha Kuznetzova)

- min degree — 0
- 4 persons have only 1 friend
- average degree ≈ 11

Number of Groups:

- max value — 355 (Danil Fyodorovykh)

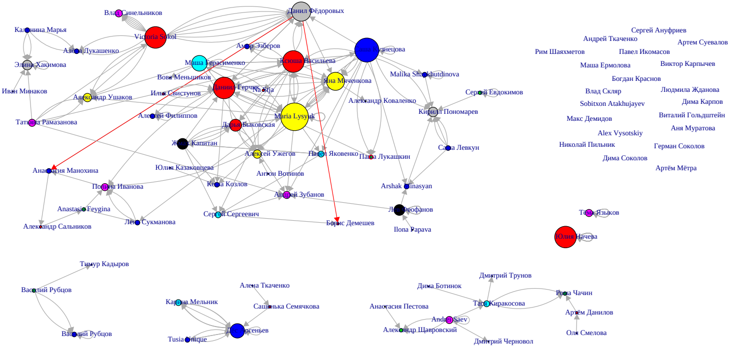
- 19 persons limit the access
- average number ≈ 43

The most popular vk groups:

1. Высшая школа экономики (24)
2. Лентач (18)
3. The Vyshka (16)

ВЫСШАЯ
ШКОЛА
ЭКОНОМИКИ

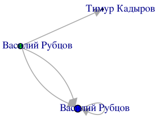
Do the likes matter?

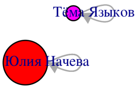


	Graph of Likes	Graph of Friends
average node degree	**2**	**11**

merely 9 last actions(photos,notes,posts) were taken into account

Do the likes matter?Some more Statistics



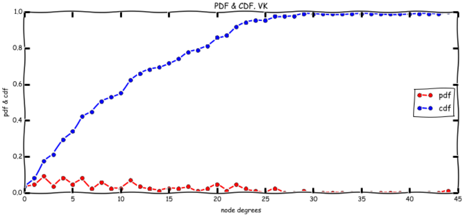


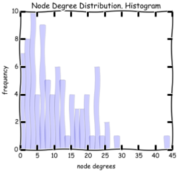
	Department of Computer Science	Department of Economic Science	total
average node degree (likes)	≈1	≈3	≈2.23
female	≈1.6	≈3.7	≈2
male	1	≈2.5	≈1.8

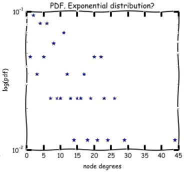
	in_likes	out_likes	efficient likes givers
1	Danil Fedorovysh (12)	Sasha Kuznetzova (18)	Danil Fedorovysh (8)
2	Yana Myachenkova (11)	Victoria Socol (17)	Kirill Ponomarev (8)
3	Maria Lysyuk (9)	Maria Lysyuk (14)	Yana Myachenkova (7)

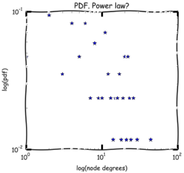
17 out of 20 persons with 0 node degree are males

Network Parameters









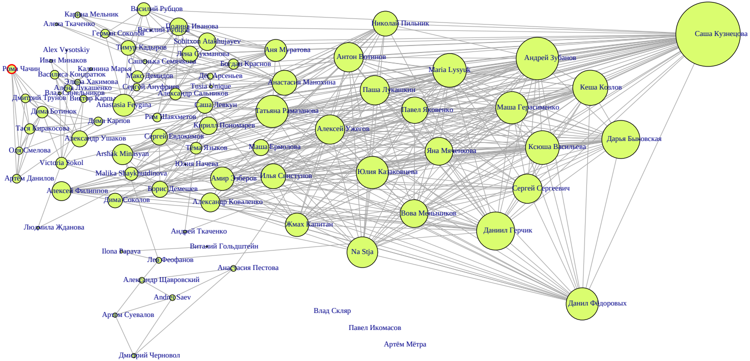
parameter	average degree	diameter	average path	average clustering
str160	float64	int32	float64	float64
value	10.66	8.00	21.25	0.59

Deduction. Part I

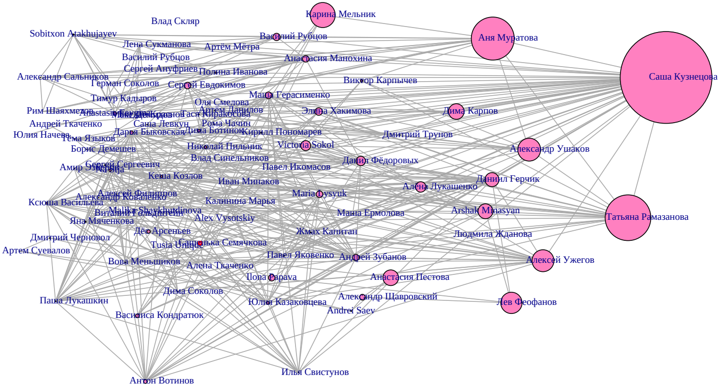
Deduction:

- *Common Friend*: a male from a economic science department of NRU HSE, who hides a relationship status and is busy enough to make posts and hand out a lot of «likes».
- *Node Degree Distribution*: while social network is suppossed to be classified as scale-free network, which means node degree distribution should follow power law, this tendency does not hold ($85 \ll \infty$).
- *Average Path*: the average path as well as the diameter of the graph is much more than 6 $\rightarrow$ the network of my vk friends does not supports the theory of 6 handshakes.
- *Cluster Coefficient*: although this coefficient should be similar to 1 for social network, it is not; there is no strong tendency for clusters distinguishing.

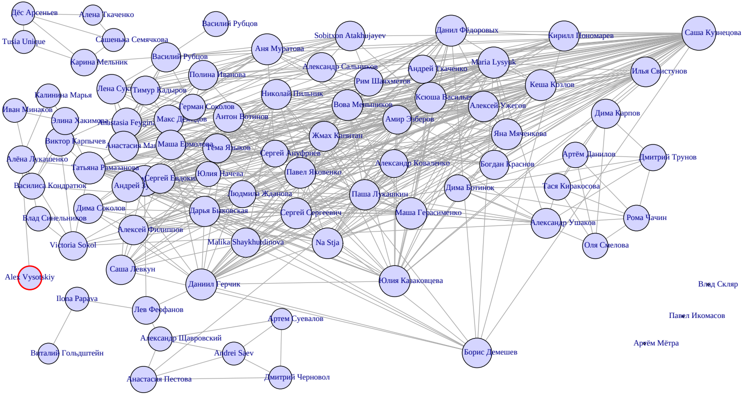
Degree Centrality



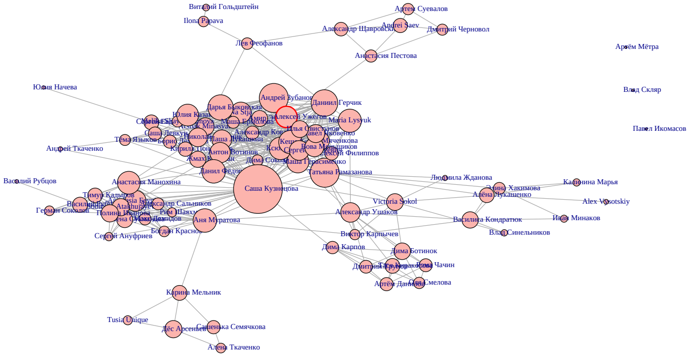
Betweenness Centrality



Closeness Centrality



Page - Rank


Виталий Гольдштейн
Ilona Papava
Лев Феофанов
Артем Суевалов
Andrei Saev
Дмитрий Черновол
Анастасия Пестова
Влад Скляр
Павел Икомасов
Юлия Начева
Андрей Зубанов
Даниил Герчик
Maria Lysyuk
Алексей Ужегов
Дарья Быковская
Маша Ермолова
Илья Свистунов
Павел Яковенко
Антон Вотинов
Маша Герасименко
Тёма Языков
Андрей Ткаченко
Василий Рубцов
Анастасия Манохина
Саша Кузнецова
Татьяна Рамазанова
Людмила Жданова
Калинина Марья
Элина Хакимова
Алёна Лукащенко
Alex Vysotskiy
Victoria Sokol
Александр Ушаков
Василиса Кондратюк
Иван Минаков
Влад Синельников
Виктор Карпычев
Дима Карпов
Дима Ботинок
Тимур Кадыров
Александр Сальников
Герман Соколов
Полина Иванова
Аня Муратова
Богдан Краснов
Сергей Ануфриев
Карина Мельник
Tusia Unique
Дёс Арсеньев
Сашенька Семячкова
Алена Ткаченко

Top Nodes Comparison

degree centrality	closeness centrality	betweenness centrality	pagerank
str576	**str576**	**str576**	**str608**
Sasha Kuznetsova	Sasha Kuznetsova	Sasha Kuznetsova	Sasha Kuznetsova
Andrey Zubanov	Tatyana Ramazanova	Tatyana Ramazanova	Tatyana Ramazanova
Darya Bykovskaya	Andrey Zubanov	Anya Muratova	Andrey Zubanov
Daniil Gerchik	Masha Gerasimenko	Karina Melnik	Daniil Gerchik
Kesha Kozlov	Maria Lysyuk	Alexander Ushakov	Darya Bykovskaya
Ksyusha Vasilyeva	Yulia Kazakovtseva	Alexey Uzhegov	Danil Fyodorovykh
Maria Lysyuk	Daniil Gerchik	Lev Feofanov	Kesha Kozlov
Masha Gerasimenko	Ksyusha Vasilyeva	Anastasia Pestova	Maria Lysyuk
Yulia Kazakovtseva	Danil Fyodorovykh	Dima Karpov	Anya Muratova
Danil Fyodorovykh	Kesha Kozlov	Arshak Minasyan	Ksyusha Vasilyeva
Tatyana Ramazanova	Anya Muratova	Daniil Gerchik	Anastasia Manokhina
Na Stja	Yana Myachenkova	Alyona Lukashenko	Masha Gerasimenko
Anton Votinov	Vova Menshikov	Victoria Sokol	Yulia Kazakovtseva
Pasha Lukashkin	Na Stja	Danil Fyodorovykh	Alexey Uzhegov
Alexey Uzhegov	Pavel Yakovenko	Elina Khakimova	Na Stja

Assortative Mixing

assortativity	degree	sex	relation	graduation	faculty_name	university_name	post_number	likes_num_net	group_num
str160	**float64**	**float64**	**float64**	**float64**	**float64**	**float64**	**float64**	**float64**	**float64**
value	0.248	-0.029	0.010	0.084	0.164	0.090	-0.021	-0.019	-0.007

Enough Evident Results:
Assortative mixing takes place according to

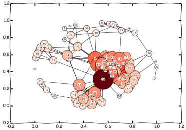
- degree
- graduation
- name of the University
- name of the Faculty

It can be interesting:
While graph is disassortative regarding to sex it at the same time assortative to relationship status
Disassortative mixing takes place according to

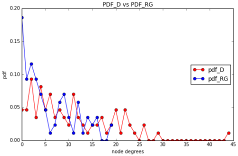
- number of posts on the wall
- amount of getting «likes»
- number of groups

The closest random graph model or how Nastya plays with parameters

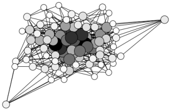
Graph of vk friends



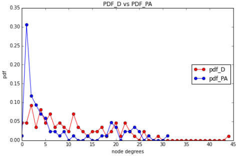
Random Graph,
$n_0 = 28, m_0 = 166, m = 5$



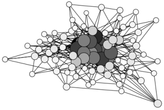
Random Graph



Preferential attachment,
$n_0 = 20, m_0 = 400, m = 4$



Preferential attachment



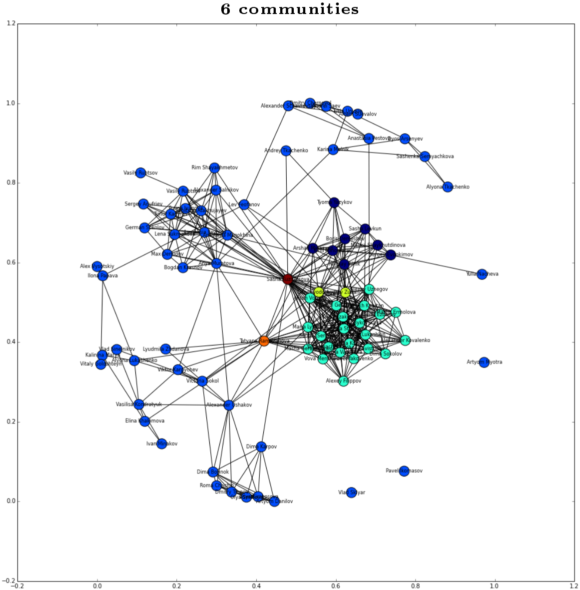
	AvND	NoEd	AvCl
Real Graph	10.66	454	0.59
RG	10.6	456	0.18
PA	10.5	446	0.31

Cliques search

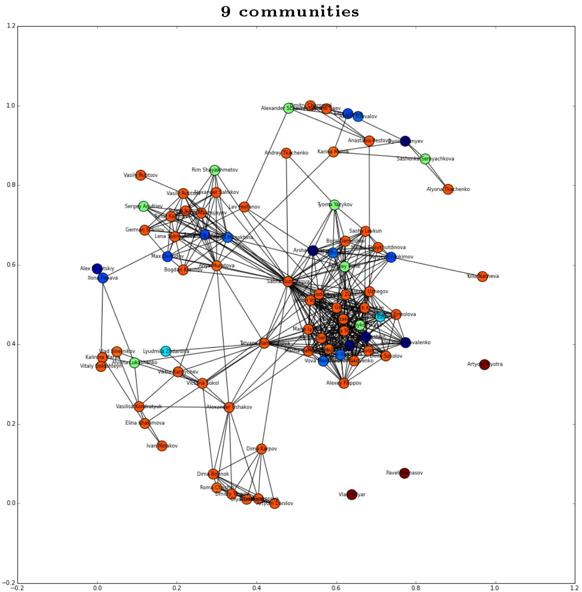
There are 179 cliques with size ranging from 3 to 12.

The Largest Clique

Hierarchical Clustering

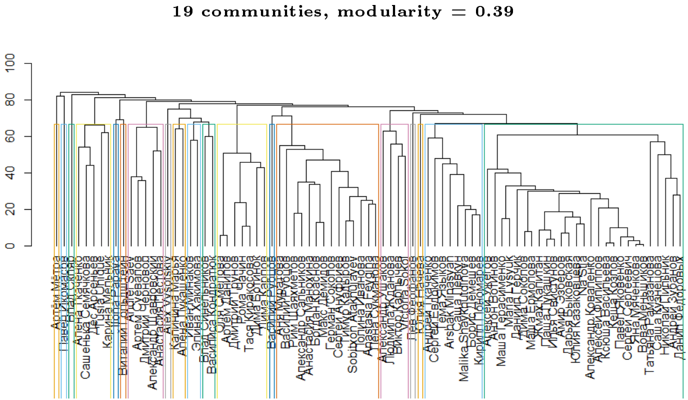


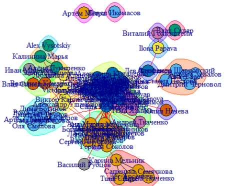
MCL



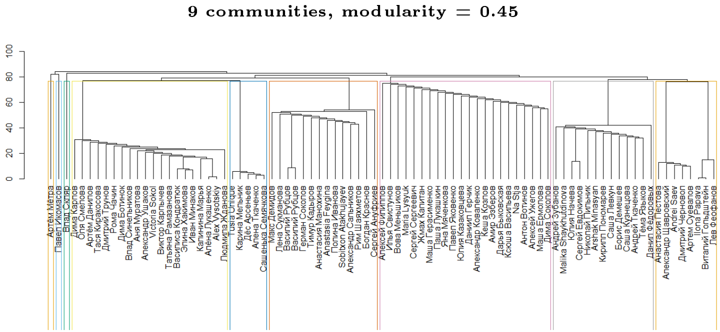
9 communities
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Vasily Rubtsov
Rim Shayakhmetov
Alexander Salnikov
Sergey Anufriev
Lev Feofanov
German Sokolov
Lena Sukmanova
Max Demidov
Bogdan Krasnov
Anya Muratova
Alex Zhitskiy
Ilona Panava
Andrey Tkachenko
Karina Melnik
Anastasia Nestova
Sashenka Semyachkova
Alyona Tkachenko
Tyoma Yazykov
Sasha Levkun
Yulia Nacheva
Sasha Kuznetsova
Tatyana Ramanova
Vlad Sinelnikov
Lyudmila Zhdanova
Alyona Lukashenko
Viktor Kanychev
Victoria Sokol
Vasilisa Kondratyuk
Alexander Ushakov
Elina Khakimova
Ivan Minakov
Dima Karpov
Dima Bobrok
Roma Chachin
Artyom Danilov
Alexey Filippov
Dima Sokolov
Artyom Ryotra
Pavel Romasov
Vlad Sulyar

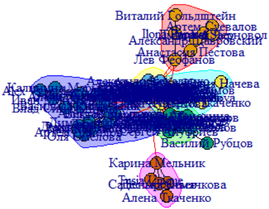
Walktrapigraph



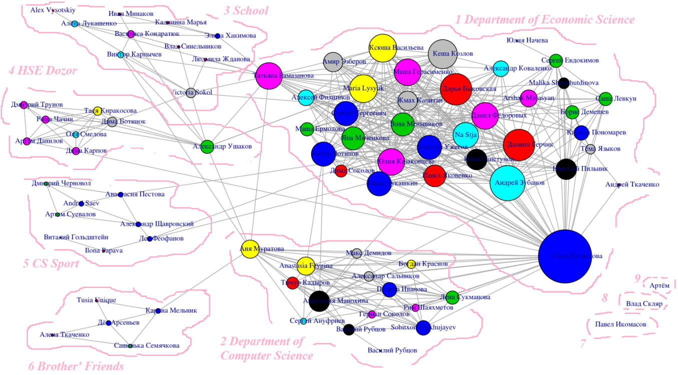


Fastgreedyigraph





True Communities



Fastgreedy Method provides the best results both in modularity and closeness to the true communities' partition terms

That is all :)

Thanks for your attention!

Words Cloud

The most frequent words occuring in posts of students studying at computer science department