1. Course Description

a. **Title:** "Social Network Analysis"

Author: Leonid E. Zhukov, National Research University Higher School of Economics, Faculty of Computer Science, Department of Data Analysis and Artificial Intelligence, professor.

Co-author: Ilya A. Makarov, National Research University Higher School of Economics, Faculty of Computer Science, Department of Data Analysis and Artificial Intelligence, senior lecturer, deputy head.

b. **Prerequisites:** The course assumes no prior knowledge of statistics and is based on basic notions of mathematics from high-school. However, all students are advised to be prepared for studying a mathematical discipline, even if they come from a non-mathematical background.

The following knowledge and competence are needed to study the discipline:

- A good command of the English language, both orally and written.
- A basic knowledge of mathematics
- A basic programming experience

c. **Course Type**: elective, MAGoLEGO course.

Place of the discipline in the MagoLego course from University pool. The course «Social Network Analysis» is a course taught in MagoLego course from University pool. It is recommended for non-specialists who wish to get fundamental knowledge in analysis of social networks

Scope of Use. The present program establishes minimum demands of students' knowledge and skills, and determines content of the course.

The present syllabus is aimed at department teaching the course, their teaching assistants, and students of MagoLego course from University pool.
This syllabus meets the standards required by:

- Educational standards of National Research University Higher School of Economics;
- Educational program of Federal Master's Degree Program for 2016 https://www.hse.ru/org/hse/elective_courses/MG_KK;
- University curriculum of the Master's program in «Data Science» (010402) for 2016.

d. **Abstract** (Summary of the course). The course presents mathematical methods and computational tools for Social Network Analysis (SNA). SNA was pioneered by sociologist, but recently became an interdisciplinary endeavor with contributions from mathematicians, computer scientists, physicists, economists etc., who brought in many new tools and techniques for network analysis. In this course we will start with basic statistical descriptions of networks, analyze network structure, roles and positions of nodes in networks, connectivity patterns and methods for community detection. In the second part of the course we will discuss processes on networks and practical methods of network visualization. We conclude the course with examples from social media mining and Facebook, Vkontakte and Twitter analysis.

2. Learning Objectives

The learning objective of the course «Social Network Analysis» is to provide students with essential knowledge of network analysis applicable to real world data, with examples from today's most popular social networks.

3. Learning outcomes

After completing the study of the discipline « Social Network Analysis » the student should:

- Know basic notation and terminology used in network science
- Be able to visualize, summarize and compare networks
- Understand basic principles behind network analysis algorithms
- Develop practical skills of network analysis in R programming language
- Be capable of analyzing real work networks

After completing the study of the discipline «Social Network Analysis» the student should have the following competences:

Competence	Code	Code (UC)	Descriptors (indicators of achievement of the result)	Educative forms and methods aimed at generation and development of the competence
The ability to reflect developed methods of activity.	SC-1	SC-M1	The student is able to reflect developed network methods in social sciences	Lectures and tutorials.
The ability to propose a model to invent and test methods and tools of professional activity	SC-2	SC-M2	The student is able to visualize and summarize data, develop mathematical models	Examples covered during the lectures and tutorials. Assignments.
Capability of development of new research methods, change of scientific and industrial profile of self-activities	SC-3	SC-M3	Students obtain necessary knowledge in network science, sufficient to develop new methods in other disciplines.	Assignments, additional material/reading provided.
The ability to describe problems and situations of professional activity in terms	PC-5	IC-M5.3_5.4_5.6_2.4.1	The student is able to describe problems in terms of network science	Lectures and tutorials.

Competence	Code	Code (UC)	Descriptors (indicators of achievement of the result)	Educative forms and methods aimed at generation and development of the competence
of humanitarian, economic and social sciences to solve problems which occur across sciences, in allied professional fields.				

4. Course Plan

Two pairs consist of 2 academic hour for lecture followed by 2 academic hour for computer exercises/labs after lecture. Additional office hours for lectures' content are provided.

№	Topic	Total hours	Contact hours		Self-study
			Lectures	Seminars	
1.	Introduction to social network analysis	10	2	2	6
2.	Descriptive network analysis	10	2	2	6
3.	Network structure	11	2	2	7
4.	Node centralities and ranking on network	11	2	2	7
5.	Network communities	11	2	2	7
6.	Affiliation networks	11	2	2	7
7.	Information and influence propagation on networks	11	2	2	7
8.	Network visualization	11	2	2	7
9.	Social media mining	11	2	2	7
10.	SNA in real world: FB/VK and Twitter analysis	11	2	2	7
Total:		**108**	**20**	**20**	**68**

Course Description

The following list describes the topics that will be covered in the course in correspondence with lecture order.

Topic 1. Introduction to social network analysis
Introduction to new science of networks. Networks examples. Graph theory basics

Topic 2. Descriptive network analysis
Statistical network properties. Degree distribution, clustering coefficient. Frequent patterns. Network motifs. Cliques and k-cores

Topic 3. Network structure.
Nodes and edges, network diameter and average path length.

Topic 4. Node centralities and ranking on network
Node centrality metrics: degree, closeness and betweenness centrality. Eigenvector centrality and PageRank. Algorithm HITS

Topic 5. Network communities
Networks communities. Graph partitioning and cut metrics. Edge betweenness. Modularity clustering.

Topic 6. Affiliation networks
Affiliation network and bipartite graphs. 1-mode projections. Recommendation systems

Topic 7. Information and influence propagation on networks
Social Diffusion. Basic cascade model. Influence maximization. Most influential nodes in network

Topic 8. Network visualization
Network visualization and graph layouts. Graph sampling. Low-dimensional projections

Topic 9. Social media mining
Natural language processing and sentiment mining.

Topic 10. SNA in real world: FB/VK and Twitter analysis
Properties of large social networks: friends, connections, likes, re-tweets

5. Reading List and Materials

We do not follow a particular textbook in this subject, but the student may find the following references useful:

5.1 Recommended Reading

1. David Easley and John Kleinberg. "Networks, Crowds, and Markets: Reasoning About a Highly Connected World." Cambridge University Press 2010.
2. Eric Kolaczyk, Gabor Csardi. “Statistical Analysis of Network Data with R (Use R!)”. Springer, 2014.
3. Stanley Wasserman and Katherine Faust. "Social Network Analysis. Methods and Applications." Cambridge University Press, 1994

5.2 Supplementary Reading

1. Maarten van Steen. “Graph Theory and Complex Networks. An Introduction”, 2010.
2. Reza Zafarani, Mohammed Ali Abbasi, Huan Liu. “Social Media Mining: An Introduction”. Cambridge University Press 2014.
3. Maksim Tsvetovat and Alexander Kouznetsov. “Social Network Analysis for Startups”. O’Reilly Media, 2011.

5.3 R programming

1. Robert Knell. “Introductory R: A Beginner's Guide to Data Visualisation, Statistical Analysis and Programming in R”, 2013
2. Alain Zuur, Elena Ieno, Erik Meesters. “A Beginner's Guide to R (Use R!)”, Springer, 2009
3. Robert Kabacoff. “R in action. Data Analysis and graphics with R”, Manning

Publications, 2011
4. Jared Lander. “R for Everyone: Advanced Analytics and Graphics”, Addison-Wesley, 2013

5.4 Popular Reading

1. Albert-Laszlo Barabasi. “Linked. The New Science of Networks”, 2002,2014
2. Duncan Watts. “The Science of a Connected Age”, 2004

5.5 Course webpage

Students are provided with links to the lecture notes, problem sheets and their solutions, assignments and their solutions, and additional readings.

6. Grading System

Type of grading	**Type of work**		
	Homework	1	Solving homework task and examples.
	Exam	1	Written exam. Preparation time – 120 min.
Final			

The assessment consists of one homework, handed out to the students during the semester. The homework problems are based on each lecture topics.

Final assessment is the final exam. Students have to demonstrate knowledge of probability and statistics theory.

7. Guidelines for Knowledge Assessment

The grade formula:

The exam will consist of 10 problems, giving ten marks each, worth 40% of the final mark.

Final course mark *is obtained from the following formula*: Final=0.6*(Homework)+0.4*(Exam).

The grades are rounded in favour of examiner/lecturer with respect to regularity of class and home works. All grades, having a fractional part greater than 0.5, are rounded up.

Table of Grade Accordance

Ten-point Grading Scale	**Five-point Grading Scale**	
1 - very bad 2 – bad 3 – no pass	Unsatisfactory - 2	**FAIL**
4 – pass 5 – highly pass	Satisfactory – 3	

6 – good 7 – very good	Good – 4	**PASS**
8 – almost excellent 9 – excellent 10 – perfect	Excellent – 5	

8. Methods of Instruction

Course lecturer is advised to use interactive learning methods, which allow participation of the majority of students, such as slide presentations, combined with writing materials on board, and usage of interdisciplinary papers to present connections between probability theory and statistics. The course is intended to be adaptive, but it is normal to differentiate tasks in a group if necessary, and direct fast learners to solve more complicated tasks.

Term Educational Technology

The following educational technologies are used in the study process:

- discussion and analysis of the results during the computer exercises;
- regular assignments to test the progress of the student;
- consultation time on Monday mornings with lecturer and after lecture;
- teleconference lectures
- office hours and classes with tutor and teaching assistants
- tutorship

Recommendations for students

The course is interactive. Lectures are combined with exercises. Students are invited to ask questions and actively participate in group discussions.

There will be special office hours for students, which would like to get more precise understanding of each topic. The first section will be devoted to learning basics of R Programming and IGraph package. The second (main) part will deal with current class and home works, helping students to understand precisely basic ideas of algorithms for SNA. The thirds and final part will go for final task evaluation and preparation for exam test for discipline.

The lecturer is ready to answer your questions online by official e-mails that you can find in the "contacts" section**.** Additional references found in section 15.1 are suggested to help students in their understanding of the material. This course is taught in English, and students can ask teaching assistants to help them with the language.

In addition to introductory classes on R language you may find useful course from Coursera: https://www.coursera.org/course/rprog

Final exam questions

The final exam will consist in ten questions equally weighted. No material is allowed for the exam. Each question will focus on a particular topic presented during the lectures.

9. Special Equipment and Software Support

The course requires a laptop and projector.
R statistical modeling environment, RStudio IDE, igraph library
R: http://www.r-project.org

RStudio: http://www.rstudio.com
Igraph: http://igraph.org

Lecture materials, course structure and the syllabus are prepared by Leonid Zhukov.

Also the syllabus part concerning classes' structure was made in collaboration with Ilya Makarov.