Рабочая программа дисциплины «Анализ сетевых сетей»
(на английском языке)
«Network Science»

для образовательной программы «Науки о данных»
направления подготовки 01.04.02. Прикладная математика и информатика
уровень - магистр

Разработчик программы
Л.Е. Жуков, PhD, профессор, ljzukov@hse.ru
И.А. Макаров, iamakarov@hse.ru

Одобрена на заседании департамента анализа данных и искусственного интеллекта
«___»___________ 2015 г.
Руководитель департамента анализа данных и искусственного интеллекта Школы
С.О. Кузнецов _________

Утверждена Академическим советом образовательной программы
«___»___________ 2015 г., № протокола_______________

Академический руководитель образовательной программы
С.О. Кузнецов ________________

Москва, 2015
Настоящая программа не может быть использована другими подразделениями университета и
другими вузами без разрешения подразделения-разработчика программы.
1. Teachers

Author, lecturer: Leonid Zhukov, National Research University Higher School of Economics, Department of Data Analysis and Artificial Intelligence, professor

Tutor: Ilya Makarov, National Research University Higher School of Economics, Department of Data Analysis and Artificial Intelligence, senior lecturer, deputy head

2. 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 the Master of Science program 010402 «Data Sciences».

This syllabus meets the standards required by:
- Educational standards of National Research University Higher School of Economics;
- Educational program «Data Science» of Federal Master’s Degree Program 010402 «Applied Mathematics and Informatics», 2015;
- University curriculum of the Master’s program in «Data Science» for 2015.

3. Summary

The course “Network Science” introduces students to new and actively evolving interdisciplinary field of network science. Started as a study of social networks by sociologists, it attracted attention of physicists, computer scientists, economists, computational biologists, linguists and others and become a truly interdisciplinary field of study. In spite of the variety of processes that form networks, and objects and relationships that serves as nodes and edges in these networks, all networks poses common statistical and structural properties. The interplay between order and disorder creates complex network structures that are the focus of the study. In the course we will consider methods of statistical and structural analysis of the networks, models of network formation and evolution and processes developing on network. Special attention will be given to the hands-on practical analysis and visualization of the real world networks using available software tools and modern programming languages and libraries.

4. Learning Objectives

Learning objectives of the course “Network Science” are to familiarize students with a new rapidly evolving field of network science, and provide practical knowledge experience in analysis of real world network data.

5. Learning outcomes

After completing the study of the discipline “Network Science”, the student should:
- Know basic notions and terminology used in network science
- Understand fundamental principles of network structure and evolution
- Learn to develop mathematical models of network processes
- Be capable of analyzing real world network data

After completing the study of the discipline “Network Science” the student should have the following competences:
<table>
<thead>
<tr>
<th>Competence</th>
<th>Code</th>
<th>Code (UC)</th>
<th>Descriptors (indicators of achievement of the result)</th>
<th>Educative forms and methods aimed at generation and development of the competence</th>
</tr>
</thead>
<tbody>
<tr>
<td>The ability to reflect developed methods of activity.</td>
<td>SC-1</td>
<td>SC-M1</td>
<td>The student is able to reflect developed mathematical methods to network sciences</td>
<td>Lectures and tutorials, group discussions, presentations, paper reviews.</td>
</tr>
<tr>
<td>The ability to propose a model to invent and test methods and tools of professional activity</td>
<td>SC-2</td>
<td>SC-M2</td>
<td>The student is able to improve and develop research methods of linear optimization, approximation and computational problem solvation.</td>
<td>Classes, home works.</td>
</tr>
<tr>
<td>Capability of development of new research methods, change of scientific and industrial profile of self-activities</td>
<td>SC-3</td>
<td>SC-M3</td>
<td>The student obtain necessary knowledge in network science, which is sufficient to develop new methods on other sciences</td>
<td>Home tasks, paper reviews,</td>
</tr>
<tr>
<td>The ability to describe problems and situations of professional activity in terms of humanitarian, economic and social sciences to solve problems which occur across sciences, in allied professional fields.</td>
<td>PC-5</td>
<td>IC-M5.3_5, 4_5.6_2. 4.1</td>
<td>The student is able to describe network problems problems in terms of computational mathematics.</td>
<td>Lectures and tutorials, group discussions, presentations, paper reviews.</td>
</tr>
<tr>
<td>The ability to detect, transmit common goals in the professional and social activities</td>
<td>PC-8</td>
<td>SPC-M3</td>
<td>The student is able to identify mathematical aspect in network research evaluate correctness of the used methods, and their applicability in each current situation</td>
<td>Discussion of paper reviews; cross discipline lectures</td>
</tr>
</tbody>
</table>
6. **Place of the discipline in the Master’s program structure**

The course “Network Science” is a course taught in the first year of the Master’s program 010402 “Data Sciences” and is a base course for specialization “Intelligent Systems and Structural Analysis”

**Prerequisites**

The course is based on knowledge and understanding of

- Discrete mathematics
- Algorithms and data structures
- Linear algebra
- Theory of probability and statistical analysis

It also requires some programming experience in one of programming languages:

- Python
- Matlab
- R

7. **Schedule**

One pair consists of 1 academic hour for lecture and 1 academic hour for classes after lecture.

<table>
<thead>
<tr>
<th>№</th>
<th>Topic</th>
<th>Total hours</th>
<th>Contact hours</th>
<th>Self-study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lectures</td>
<td>Seminars</td>
</tr>
<tr>
<td>1</td>
<td>Introduction to network science</td>
<td>9</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Power laws</td>
<td>13</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Models of network formation</td>
<td>15</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Structure, nodes and links analysis</td>
<td>17</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Network communities</td>
<td>17</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>Diffusion and epidemics on networks</td>
<td>15</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Influence propagation</td>
<td>17</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>Information cascades</td>
<td>15</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>Evolving networks and link prediction</td>
<td>17</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>Network visualization</td>
<td>17</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>152</strong></td>
<td><strong>30</strong></td>
<td><strong>46</strong></td>
</tr>
</tbody>
</table>
8. Requirements and Grading

<table>
<thead>
<tr>
<th>Type of grading</th>
<th>Type of work</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Homework</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Special homework – research projects</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mid Term Exam</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Final Exam</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

9. Assessment

The assessment consists of weekly classwork and homework. Short quizzes are assigned after each lecture through the google-forms to make students pass their homeworks (otherwise the maximal grade for HW will not exceed 4 of 10). Students have to demonstrate their knowledge in each lecture topic concerning both theoretical facts, and practical tasks’ solving. All tasks are connected through the discipline and have increasing complexity. The two practical projects should be presented in class and graded by both, lecturers and students. The weight of a project is three times greater than the weight of each HW.

Each technical report should be made in IPython Notebook format. Labs are to be uploaded in DataJoy Group storage.

Students obtain the following points per each condition:
- Correct data analysis – 4 points;
- Suitable visualization – 3 points;
- Report completeness and formulas’ descriptions – 2 points;
- Extra 1 point for overall report evaluation and excellent research and soft skills.

Technical details for reports and projects are placed to the corresponding technical section.

Final assessment is the final exam. Students have to demonstrate knowledge of theory facts, but the most of tasks would evaluate their ability to solve practical examples, present straight operation, and recognition skills to solve them.

The grade formula:

The exam will consist of 10 problems, giving 10 points each, total 100 points for the exam

Final course mark is obtained from the following formula:

\[ O_{\text{final}} = 0.6 \times O_{\text{cumulative}} + 0.4 \times O_{\text{exam}}. \]

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.
10. Course Description

The following list describes main topics covered by the course with lecture order.

**Topic 1. Introduction to network science**

**Content:**

**Recommended reading:**

**Supplementary reading:**

**Topic 2. Power laws**

**Content:**

**Recommended reading:**
1. Chapter 8 of Mark Newman. "Networks: An Introduction". Oxford University Press,
2010.


Supplementary reading:

Topic 3. Models of network formation

Content:

Recommended reading:

Supplementary reading:

Topic 4. Structure, nodes and links analysis

Content:
Recommended reading:

Supplementary reading:
3. S. Brin, L. Page. The PageRank Citation Ranknig: Bringing Order to the Web.

Topic 5. Network communities

Content:

Recommended reading:

Supplementary reading:
Topic 6. Diffusion and epidemics on networks

Content:

Recommended reading:

Supplementary reading:

Topic 7. Influence propagation

Content:

Recommended reading:

Supplementary reading:

**Topic 8. Information Cascades**

**Content:**
Observational learning. Information cascades.

**Recommended reading:**

**Supplementary reading:**
4. S. Bikhchandani, D Hirshleifer and I.Welch. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades

**Topic 9. Evolving networks and link prediction**

**Content:**
Recommended reading:

Supplementary reading:

Topic 10. Network visualization Content:

Recommended reading:

Supplementary reading:

11. Term Educational Technology
The following educational technologies are used in the study process:
• discussion and analysis of the results of the home task in the group;
• individual education methods, which depend on the progress of each student;
• analysis of skills to formulate common problem in terms of mathematics and solve it;
The course is used to be in format of teleconference lectures.
12. Recommendations for course lecturer

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 interactive programming environments for demonstration purposes. 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.

13. Recommendations for students

The course is interactive. Lectures are combined with classes. 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. Teaching assistant will also help you. All tutors are ready to answer your questions online by official e-mails that you can find in the “contacts” section.

14. Sample final exam questions

1. Some social network has an exponential distribution of node degrees, given by $P(k) = Ce^{-ak}$, where $C$ and $a$ are constants and $a$ is given. Find the fraction of the network nodes that have no more than $k_{\text{min}}$ neigbours. Use continues approximation $k_{\text{min}} = 0$, $k_{\text{max}} = \infty$.

2. Consider random walk on the undirected graph given below. Calculate the probability of being on the node #9 in the limit of large time.

3. Describe Zipf’s law. What is the connection between Zipf’s law and power law distributions?

4. Calculate node degree distributions in the Barabasi-Albert model with uniform probability of attachment $P(k) = 1/(m_0 + T)$

5. Which of the graphs given below has been most likely generated form Erdos-Renyi model with $N=5$ and $p=0.3$ Why? Which one couldn’t be generated by that model?
15. Reading and Materials

15.1 Required Reading


15.2. Recommended Reading


15.3. List of review papers


15.4. Course telemaintenance

All material of the discipline are posted in informational educational site at NRU HSE portal www.cs.hse.ru/ai, http://www.leonidzhukov.net/hse/2016/networks/, https://github.com/MakarovIA/networks and DataJoy Group. Students are provided with links to research papers, electronic books, data and software via different cloud services and tools.

16. Equipment

The course requires a laptop, projector, and acoustic systems. It also requires opportunity to install programming software, such as:

- Python
- Matlab
- R

on student personal computers. One of core ways is to use the service www.getdatajoy.com or use installed software on
Amazon Web-server.

Lecture materials, course structure and syllabus are prepared by Leonid Zhukov and Ilya Makarov