Skoltech
Skolkovo Institute of Science and Technology

Thesis Changes Log

Name of Candidate: Saeed Osat

PhD Program: Computational and Data Science and Engineering

Title of Thesis: Percolation on complex networks and its applications

Supervisor: Assistant Professor Vladimir Palyulin

The thesis document includes the following changes in answer to the external review process.

- **page 4.** In the 3rd column in the table the definition is updated.
- **Page 5.** Definition of "Excess degree" and "Excess degree distribution" is added to the table.
- **Page 15.** The whole section 1.1.11 is added in the revision, describing phase, phase transition and its types.
- **Page 18.** The whole section 1.1.12 is added in the revision, describing the method of generating function.
- **Page 22.** Definition of susceptibility and its relevance to transition is added.
- **Page 24.** Theoretical prediction for the phase transition of GMCC is added.
- **Page 27.** Caption of the Figure 1.4 is updated
- **page 28.** The analogy between ICM and BP is explained.
- **Page 35.** P_∞ for observability model is defined.

Point-to-point address of the jury reports

Saeed Osat *October 11, 2021*

Here is the point-to-point response to the reports of the jury members. The gray boxes contain the reports and the red boxes the reply of the student. All the changes made in the final version of the thesis are listed in the **logfile**.

Prof. Brilliantov's Report

The work "PERCOLATION ON COMPLEX NETWORKS AND ITS APPLICATIONS" submitted as a PhD thesis by Saeed Osat studies an important novel type of complex networks called multiplex networks. The thesis is submitted in a form of a coherent academic treatise consists of 8 chapters and is based on 7 papers published in scientific journals. Among them there are 3 first author papers and 4 second author papers. 5 articles are published in Q1/Q2 journals indexed in WOS/Scopus with 2 among them counted in the Nature Index. This securely fulfills the requirements of the PhD thesis defense policy of Skoltech. The first out of the 8 chapters is an introduction. The following 6 chapters consist each of a text of the corresponding paper and a short description of the contribution of the applicant to the published work. The final chapter summarizes the scientific contributions of the thesis. Overall, I am very satisfied by the scope and the comprehensiveness of the study of the relatively new class of complex networks. The model of multiplex networks is a very useful approach for the description of complex transportation systems, www networks, social networks and others. The practical importance of obtained results is clearly shown in the text. The treatise presents a significant advancement in the field. The high-quality papers support this statement and show the author's ability to utilize state-of-the-art computational methods as well as deep knowledge of the relevant literature. The first part of the introduction covers the basics of the classical network theory as well the generalization of definitions and notions such as k-core, percolation, giant component, community structure and geometric correlations. It also introduces the new notions which are specific for the multiplex case including interlayer degree correlation and the edge overlap. The second part of the chapter describes the contents of the papers attached together as a treatise and binds them together in a cohesive form. The second chapter considers a simple model of neuronal dynamics on networks to show that the degree distribution of the underlying neuronal network may trigger power-law distributions for neuronal avalanches. Contrary to the previous opinion in the literature, this could happen even in the case when the system is not in a critical regime. The third chapter generalizes the notion of k-core used in simple networks to the multiplex networks and is called Gk-core (generalized k-core). The analytical and numerical approaches for Gk-core percolation were applied to real-world networks such as transcription networks, yeast-protein networks, computer networks etc. The fourth chapter is devoted to observability transition and generalization of the notion of observability to multiplex networks. Again, the approach was tested on the real-world networks (US air transportation multiplex network, C. Elegans connectome, Arxiv, etc.). Chapter five addresses the embedding of multiplex networks into a hyperbolic space. The correspondence of community structure and hyperbolic embedding is shown. Then the correlations between community structure between the layers and the similarity of the nodes in the embedding is revealed. Chapter six culminates in consideration of a very physically important question, namely, of an optimal percolation in multiplex networks. This part of work pioneers this question in the field of multiplex networks and discovers its essential difference from a simple network case. The main conclusion states that an attempt of reducing of a multiplex network to a monoplex case leads to the substantial error in identification of the set of structural nodes, i.e. neglecting of the multiplex structure may effect in a significant inaccuracy in the estimation of its robustness. The currently final chapter with the results, chapter 7, deals with the k-core structure of real multiplex networks (internet, ArXiv coauthorship network etc.). The main outcome is that for the cases with heterogeneous degree distribution of the network a strong k-core structure is well predicted by positive degree-degree correlations. In the other case of homogeneous networks the strong k-core structure appears due to positive correlations at the level of node similarities. While the general impression from the thesis is excellent there are a few questions/remarks to be mentioned. Some of them stem from my curiosity and the interest to the problem and may be rather general.

I thank prof. Brilliantov for his detailed report. Below you can find my detailed responses to each of the comments.

1. Although the authors mentioned a few times Hamiltonian of a network he does not explain its meaning in the current context. I believe that it would be worth to add somewhere a discussion about this notion. The related question – does there exist any relation/similarity between the cost function and the partition function? I believe that as far as the Hamiltonian may be introduced (as well as effective temperature) in the context of networks, one can define the partition function. Has this been done?

Regarding Hamiltonian, I was referring to the cost function of the given optimization problem. One can assume an optimization problem on top of a network i.g. decycling or dismantling problem and then encode the corresponding cost function as the Hamiltonian of the system. By treating the Hamiltonian as the Hamiltonian of an imaginary physical system, an effective temperature can be defined. One can obtain the best solution of the optimization problem in the low-temperature limit of this Hamiltonian.

2. General question: You did not discuss the notion of entropy of a network, does it make sense? Will the entropy of a multiplex network be additive with respect to the partial entropies of the layers?

There is not a generally accepted approach towards the entropy of the network. Regarding the question about the entropy of a multiplex network, there are several approaches that you can define Von Neumann entropy based on Laplacian matrix. The recently introduced definition is proved to be sub-additive for aggregation of layers of a multiplex network. (see also "PHYSICAL REVIEW X6,041062 (2016)")

3. In the case of percolation in a multiplex network can two (or more) percolating clusters coexist there? It is believed that this is not possible in the case of a common (one-layer) network.

In simple multiplex networks (the ones without specific correlations mentioned in this thesis) the percolation phase diagram is usually discontinuous and only a single giant mutually connected component (GMCC) exists. However, as in the case of explosive percolation if you define some new rules regarding the relation of the nodes in two layers you may end up in coexisting GMCC. However, the probability of coexisting GCC should go to zero in the thermodynamic limit.

4. Is it possible that a giant percolating cluster is comprised partly of nodes of a first layer and partly of nodes of the second layer of a multiplex network, while there are no percolating clusters separately in either layers?

   There are different definitions for cluster in multi-layer networks. For a specific case of a multiplex network it is assumed that each node exist simultaneously in both layers and they are just a replica of the same entity. The GMCC then is defined as set of nodes that they are mutually accessible in both layers through inclusive paths (nodes along these paths also are included in the set). This definition rules out the case of a GMCC as union of mutually exclusive components in two layers.

5. The last of the published papers on the list is neither a part of the introductory chapter nor it was added as a separate chapter at the end. I suggest to fix this by adding the corresponding parts.

   The last paper is done during my Ph.D. in collaboration with the group of Prof. Dylov and Prof. Biamonte and is not in the direction of thesis. We decided not to include it in the chapter 1 since it will reduce the coherence of the text.

6. Apart from this I have a list of minor corrections:

   - There are a few typos across the manuscript to be taken care of.
   - Everywhere across the text: The shortening "Fig" should be followed by the full stop, i.e. "Fig."
   - Page 4 in the definition of the clustering coefficient there should be a comma after the formula and before the word "where"
   - Page 43. The sentence "As clear from the figure, randomization weakens the k-core spectrum." I would recommend to specify the number of the figure and add a definite article in from of "randomization"

- Page 47. “Writing of the manuscript was mostly done by the main Author Ali Faqeeh. All the authors contributed in the editing of the paper.” The word author in the first sentence should be decapitalized and “contributed in” changed to “contributed to”

All the suggested corrections are applied.

Prof. Kitsak's Report

Report on the Ph.D. thesis titled "PERCOLATION ON COMPLEX NETWORKS AND ITS APPLICATIONS" by Saeed Osat. The present thesis is based on seven original research works that Saeed published during his Ph.D. studies. The thesis discusses the percolation properties of real and synthetic networks and their applications to neuronal avalanches, epidemic spreading, diffusion, and routing problems. The thesis is well-written and is in the coherent academic treatise format. The thesis' structure is adequately summarized in the Table of Contents, page vi of the document. For brevity, I refrain from typing the detailed summary of the thesis and instead focus on methods, results, and the impact of Saeed's works that are presented herein. It is well-accepted that big data often comes in the form of networks or graphs. One example is a social network, where individuals are abstracted as nodes or vertices, and interactions between individuals are abstracted as links or edges. A relatively new domain in network and data sciences is that of multiplex networks. Interactions comprising a network can be of different types. A colloquial example, to this end, is a transportation network where nodes are regions of interest and links correspond to different transportation modes (bus, train, airline, personal auto) connecting them. The core of Saeed's Ph.D. work is the development of novel percolation techniques for the analysis of large multiplex networked systems. Historically, the term percolation originated in Polymer physics, where it was introduced to study the properties of porous materials. Percolation has been reinvented in studies of large networks, where it has been extensively used in (cyber)security, epidemic spreading, network control, and communication applications. In most general terms, network percolation refers to the analysis of the network under the removal or addition of its elements (nodes or links). In this Ph.D. research, Saeed has proposed novel percolation techniques and applied them to multiplex networks. K-core percolation has been found to be an instrumental tool in a variety of applications, ranging from data visualization to identification of influential spreaders. In his work published in Physical Review E (publication 4, page iv and Chapter 3 of the thesis), Saeed extended the notion of k-core percolation to Gk-core and developed mathematical models to study them in random mono-layer networks. Gk-core can be viewed as the generalized measure quantifying the robustness of networks to various types of adverse events. In his related work (publication 6, page iv, and Chapter 7 of the thesis), Saeed has extended the notion of k-core percolation from conventional to multiplex networks. His theoretical and computational methods developed in this work set the foundation for the mathematical analysis of multiplex networks. In particular, Saeed was able to explain the roles of node connectivity patterns and correlations between network layers on the resulting structure of k-cores. These results, published in Physical Review Research, are of great importance for epidemic spreading and peer influence studies. Of central importance to both epidemic studies and studies of social influence is the optimal attack problem, where one is interested in finding the minimal set of nodes or links, that if removed, will lead to the collapse of a network into fragmented pieces. The optimal attack problem is of direct relevance to the optimal vaccination/immunization problem, where one is interested in vaccinating the population with limited resources to achieve the maximal suppression of the viral process. In a seminal work published in the prestigious Nature Communications journal (publication 1 page iv and Chapter 6 of the thesis), Saeed has developed the generalization of optimal percolation problem to multiplex systems, demonstrating that "multiplex networks have considerably smaller sets optimal sets compared to their single-layer based network representations. Arguably, the two state-of-the-art models of real networks are the Stochastic Block Model (SBM) and the Random Hyperbolic Graph (RHG). While introduced with different applications in mind, both the SBM and the RHG models have become popular modeling tools due to their ability to reproduce complex features of real networks, including heterogeneous degree distributions, topological clustering, and community structure. In his recent work (publication 3, page iv and Chapter 5 of the thesis), Saeed demonstrates that there is a deep connection between the two. His results, published in Physical Review Letters, demonstrate that the "community structure of a network can be used as a coarse version of its embedding in a hidden space with hyperbolic geometry." This finding is extremely important as it connects two challenging data science problems: community finding and geometric representation of large networks. Network observability is the problem of measuring network state using a minimal number of sensors. In his Physica A publication (publication 2, page iv and Chapter 4 of the thesis), Saeed managed to extend the notion of network observability to 2-layer multiplex networks and developed a mathematical framework for the analysis of the macroscopic cluster of mutually observable nodes. These findings are of direct relevance to the control theory of networked systems. While equally important in practice, negative results are often under-reported compared to positive results due to the perceived lack of impact. One of Saeed's works (publication 5, page iv and Chapter 2 of the thesis) can be regarded as a negative result. One of the common beliefs is that the appearance of power-law avalanches in the dynamics of neuronal networks is the signature of criticality: neuronal networks are believed to exist in a "critical" state, in between a phase where activity rapidly dies out and a phase where activity is amplified over time. Together with collaborators, Saeed has disproved this claim by introducing a counter-example. In a simple avalanche model, Saeed demonstrated that power-law distributions for neuronal avalanches might emerge even when the system is not in its critical regime. Last but not least is Saeed's study of super-diffusion in multiplex networks (publication 5, page iv of the thesis). This work demonstrates that Machine Learning can classify and predict the presence of super-diffusion in multiplex networks with high accuracy. In relation to Saeed's thesis, I only have several clarifying optional remarks that in no way diminish the quality of Saeed's work. I do, nevertheless, believe that addressing my comments might improve the accessibility of the thesis to non-specialists.

I thank Prof. Kitsak for his detailed report. Below you can find the detailed answer to each suggested corrections or additions.

1. On page 4, k-core is defined as "A maximal subset of nodes where each node has at least degree k." An inexperienced reader may interpret k-core as a set of nodes with degrees larger or equal than k. To avoid such misconception, I recommend adding "at least degree k within the subset."

   The suggestion is applied.

2. Please define phases, phase transitions, and their types before discussing "Percolation on monoplex networks," page 14

   Thank you for the suggestion. I added a new section called "Phase transition" describing the basics of phase transition and different types of it.

3. Likewise, I recommend creating a dedicated section discussing the basic formalism of generating functions in application to computing network's connected component.

   Thank you for this suggestion. A new subsection is dedicated to the basic formalism of generating fuction and its application in computing connected components of a graph.

4. The thesis routinely operates with the notion of the Giant Connected Component (GCC). GCCs are often confused with Largest Connected Components (LCC)s. To avoid this, I recommend carefully defining both terms in the thesis.

   The suggestion is applied in the new section "Phase Transition".

5. Please include the definition of the excess degree in Table 1.1

   Definition of the "excess degree" and "excess degree distribution" are added to the table.

6. Please define the susceptibility for Fig. 1.2(c). Explain the role of susceptibility in phase transition studies on complex networks, page 17.

   The suggestion is applied.

7. Clarify the "theoretical prediction" for Fig. 1.3

   The suggestion is applied.

8. Add a 1-2 sentence clarification for the analogy between the ICM and bond percolation.

   The suggestion is applied in the relevant section.

9. Consider adding a dedicated section, defining ordinary percolation, k-core percolation, core percolation, and Gk-core percolation along with the algorithms used to implement them on networks.

   The current section 1.2.2 includes definitions of K-core, core and Gk-core and their implementation algorithms.

10. Network observability needs to be defined prior to page 27. What are the different phases for observability? Please clarify on Page 27. Related to It, define the quantity plotted in Fig. 1.8

    Definition of the network observability now is added in the "Phase transition" section. P_∞ is defined in the relevant section.

In summary, Saeed published seven works in top physics and interdisciplinary journals, resulting in more than 100 citations per Google Scholar in a period of fewer than five years. His research impact is significantly higher than expected at his academic age. I would like to emphasize Saeed's communication and collaboration skills, which are remarkable for his academic age: his publications are the results of 3 distinct long-distance scientific collaborations with teams of top experts in network science from Cyprus, Portugal, and the United States. I am fully satisfied with the quality of his research. I do recommend his thesis for the defense.

Prof. Koshev's Report

The Thesis is devoted to the mathematical description of the processes occurring in complex networks. Namely, the percolation in monoplex and multiplex networks is considered. Researches in this area lead to the improvement of mathematical modeling in complex systems, and the actuality of such researches follows by both growing number of applications (neuronal networks, social networks, transport systems, etc.) and permanently improving the power of modern computational systems. The thesis starts with the introduction acquainting the reader with basics (definitions) of network analysis, a review of problems and results in this area, and a short summary of the results obtained by the Candidate. The aforementioned introduction is written in a good academic manner and meets high standards of structure, readability (including the English language), and overall quality. It avoids any ambiguity and leaves no questions, even for a reader, who is new to this area. Further chapters of the Thesis are compiled with the articles published by the Candidate during his Ph.D. study. This is possible since the Candidate published more than five articles in WoS/Scopus indexed journals, and therefore meets requirements for the representation of a Thesis as a Coherent Academic Treatise (Skoltech Ph.D. Theses Defense Policy). The structure of the Thesis also meets the requirements of the Policy. The results of research work are dedicated to both theory and applications of the percolation theory: avalanches in neuronal systems, generalization of leaf removal algorithm, an extension of the observability model to multiplex networks, optimal percolation on multiplex networks, etc. The novelty, the relevance of the topic and methods, and overall scientific significance of the aforementioned results are undoubtedly confirmed by the high level of journals publishing them: Physical Review Letters, Nature Communications. During the Ph.D. study, the Candidate published 7 articles in high-rated journals (Q1 by WoS/Scopus, nature indexed). Undoubtedly, the Candidate proved his ability to perform research meeting the highest standards and, therefore, to achieve high-rated scientific results. I definitely recommend the thesis for the presentation with the aim of receiving the Ph.D. degree. Since the Thesis is prepared in form of a coherent academic treatise and compiled of articles, the only minor comment on the thesis refers to the first – introductory – Chapter. Figure 1.4 on Page 22 should be changed since it is impossible to understand colors at the legends.

I thank Prof. Koshev for the report. The caption of the Figure 1.4 now is changed and any possible ambiguity is lifted.

Prof. Krapivsky's Report

The thesis work of Saeed Osad is devoted to the study of percolation on complex networks, observability of transitions in multiplex networks, core percolation, and kcore structure of multiplex networks. The thesis starts with a short outline and a list of publications and continues (Chap. 1) with an excellent general introduction to the science of complex networks, percolation, the notions of core percolation, and the kcore structure of networks. In this chapter, Saeed Osad also describes the main results of the published papers. The following Chapters 2–7 are the copies of the published papers. Chapter 8 gives brief conclusions. The introductory chapter 1 introduces the basic objects and concepts. Some are classical in graph theory, others are recent (multiplex networks, etc.). One beautiful concept which is almost classical, but somehow little known, is the concept of kcore and the related concept of the core. I've never heard about the latter although it is classical going back to the work of Karp and Sipser on the pruning procedure called greedy leaf removal. The giant core in a random graph emerges substantially later than the percolation transition; both transitions are continuous. An interesting generalization of the pruning transition gives birth to a giant Gk-core. All these phase transitions are very nice, and this part of the work is particularly interesting in my eye. (Admittedly, as it is mostly scientifically intriguing and perhaps a bit can be done analytically.) As in most work on networks, Erdős and Rényi are the first scientists mentioned. They wrote outstanding papers that generated great interest, but it is unfair to ignore e.g. Paul Flory who was a true pioneer and had tons of results (the cluster size distribution, the understanding of the percolation transition, etc.) much earlier than anyone else. His work was phrased using terminology from polymer science and it may be difficult to see, especially to mathematicians and computer scientists, that he was doing essentially the same. Anyway, here are a few relevant studies: P. J. Flory, J. Amer. Chem. Soc. 63, 3083 (1941); W. H. Stockmayer, J. Chem. Phys. 11, 45 (1943); P.J.Flory, Principles of Polymer Chemistry (Cornell University Press, Ithaca, 1953). I am mentioning this not just to be fair but also since a kinetic approach used in that work is overwhelmed by the probabilistic approach to random graphs frequented by mathematicians and computer scientists. Also, the original paper by Erdős and Rényi was called "Evolution of random graphs", so they thought about the process dynamically. The percolation transition in a random graph is much easier to understand, at least in my view, using the kinetic approach. So I wonder would it be possible to advance the understanding of the phase transitions related to the birth of the giant core, and more generally the giant Gkcore, using the kinetic approach? I am sure it is not a straightforward generalization of the description of the cluster size distribution, and the emergence of the giant component, etc. in the evolving random graph. So it is in principle feasible that the kinetic approach cannot handle such quantities, but it is worth trying. Another thought that has come to my mind is the concept of onion decomposition. This recent concept is very much related to the kcore decomposition. Thinking about it one starts to appreciate not only the importance of phase transitions that occur when the giant kcore, the giant Gkcore, etc. emerge in a random graph, but the concepts of kshells, the onion decomposition, and the onion layers. All those quantities can be defined for an arbitrary graph. It seems that Saeed Osad is well equipped for extending the tools he used in the PhD work to those other related and intriguing quantities. Overall, the thesis by Saeed Osad reports high-quality research. The author demonstrated originality and creativity, he established intriguing results using advanced numerical techniques, and he demonstrated theoretical strength and excellent knowledge of various branches of network science. I do not personally know Saeed Osad, but from the thesis and published papers he appears to be a mature and talented scientist and there is no question in my mind that the defense of this thesis could be made in Skolkovo Institute of Science and Technology.

I thank Prof. Krapivsky for his detailed report and comments. There are already kinetic approaches for the phase transition of cores and k-cores (chapter three and references therein). The approach mainly is based on writing the rate equation for the degree distribution. Time evolution is associated with the removal of the edges attached to k-leaves. At each time step all the edges of a randomly chosen k-leaf and its k immediate neighbors are removed. This edge removal process keeps the number of the nodes constant but changes the degrees of the nodes and thus it is necessary to do the book-keeping of the degree distribution during the process. The final equation for the evolution of the degree distribution in the case of Gk-core does not accept a closed-form solution. However, by setting a cut-off degree and solving the set of equations up to that cut-off degree one can estimate the number of nodes and links in Gk-core.

Prof. Nechaev's Report

The work of Saeed Osat is devoted to the investigation of the question how the collective statistical behavior in complex distributed systems (in particular, multiplex networks) emerges under changing the control parameters. The general unified idea lying behind all works constituting the core of the Thesis is to probe the multiplex networks for the percolation transition. I should honestly say that reading the Thesis I was pleasantly surprised by the solidness of posed questions, used tools and diversity of considered models. Saeed demonstrated very good possession of the methods of generating functions combined with exhaustive numerical investigations. The Thesis of Saeed Osat consists of 6 original chapters (one chapter per one particular posed problem, arranged as a published paper) which are preceded by the introduction serving as a navigator through the Thesis. Each chapter begins with a clear explication of the personal contribution of Saeed Osat to the work. These chapters are: 1) "Emergence of power laws in noncritical neuronal systems", 2) "Generalization of core percolation on complex networks", 3) "Observability transition in multiplex networks", 4) "Characterizing the Analogy Between Hyperbolic Embedding and Community Structure of Complex Networks", 5) "Optimal percolation on multiplex networks", 6) "k-core structure of real multiplex networks". Despite all sections of the Thesis are important and emphasize different facets of percolation on complex networks, I would like to highlight the sections 1), 4), 5) and 6), which provide elegant solutions of interesting and new questions. Let me summarize briefly the findings of these works. In the chapter "Emergence of power laws in noncritical neuronal systems" authors have demonstrated that the degree distribution of the network underlying neural dynamics plays a fundamental role in the emergence of power-law distributions of avalanches ("cascades") in their sizes. In the model under consideration, an avalanche starts with a single activated neuron and, at each time step, every one of the active neurons fires a signal that stimulates all of their neighbors. The avalanche of activities continues until no new neuron can be activated. This model is identical to the so-called independent cascade model, often considered in the context of opinion spreading in social networks. Considering a simplified model of neural dynamics on networks, authors have shown that, for some scale- free networks, avalanche sizes obey power-law distributions even in subcritical dynamical regime. In the chapter "Characterizing the Analogy Between Hyperbolic Embedding and Community Structure of Complex Networks" it is shown that the community structure of a network under embedding into the hyperbolic disc is manifested in the "closeness" of angular coordinates. Authors performed a systematic analysis of several real world and synthetic networks and have demonstrated that networks which are "natural" for hyperbolic embedding (i.e. the loss function of embedding of the network into the hyperbolic disc is relatively small) typically are highly modular, in the sense that partitions found by community detection algorithms correspond to very large values of the modularity function and nodes within the same communities are likely to have similar angular coordinates. In the chapter "Optimal percolation on multiplex networks" authors consider the problem of finding the minimal set of network nodes removal of which leads to the defragmentation of the multilayer network into non-extensive disconnected clusters. The solution to this problem is important for some practical questions, such as strategies of immunization in disease spreading, and influence maximization in opinion dynamics. Authors state that despite optimal percolation has received considerable attention in the context of isolated networks, however its generalization to multiplex networks has not yet been considered. The study is focused mostly on the characterization of the sets of "structured nodes" of a given multiplex network whose removal is crucial for the network defragmentation. It is shown that generally, multiplex networks have considerably smaller sets of structural nodes compared to the one of single-layer networks. Overall, the authors conclude that neglecting the multiplex structure may result in significant inaccuracy of predictions about the network robustness. In the chapter "k-core structure of real multiplex networks" the authors state that multiplex networks are convenient mathematical representations for many real-world systems: biological, social, and technological systems of interacting elements, where pairwise interactions among elements have different flavors. Previous studies pointed out that real-world multiplex networks display significant interlayer correlations (degree-degree correlation, edge overlap, node similarities) which are able to make them robust against random and targeted failures of their individual components. Authors show that interlayer correlations are important also for the characterization of their k-core structure, namely, for the shell organization of nodes with an increasingly high degree. Understanding the k-core structures is important in the study of spreading processes on networks – for example for the identification of influential spreaders and emergence of localization phenomena. Authors made an important conclusion that if the degree distribution of the network is heterogeneous, then a strong k-core structure is well predicted by significant positive degree-degree correlations, while if the network degree distribution is homogeneous, then strong k-core structure is due to positive correlations at the level of node similarities. Overall, the Thesis of Saeed Osat provides diversity of important new results in the field of statistical physics of complex networks. It presents a number of interconnected questions which constitute the self-contained subarea in the sea of problems related to multiplex network statistics. The thesis is clearly written and highlights practical questions beyond the purely academic framework. Minor corrections, mainly concerning the style of presentation, do not influence my very high appreciation of the work. Without any doubts, Saeed Osat deserves the Degree of "Doctor of Philosophy" with the grade magna cum laude (very good) and his Thesis satisfies all relevant requirements.

I thank Prof. Nechaev for his report.

Prof. Panov's Report

Saeed's thesis is devoted to the percolation on complex networks and its applications. In particular, he starts from studying different types of percolation for monoplex networks and then generalizes the results to the case of multiplex networks. Special attention is paid to community structure in networks and its relation to hyperbolic embeddings. The thesis is well written and follows the conventional structure: introduction, series of chapters devoted to methodology accompanied with experimental results, and conclusions. The main content is concentrated in the following chapters with their titles well corresponding to the content: emergence of power laws in noncritical neuronal systems, generalization of core percolation on complex networks, observability transition in multiplex networks, optimal percolation on multiplex networks and k-core structure of real multiplex networks. Each of these five chapters is followed by a discussion on the presented results. From the complex networks perspective, relevant methods are used. The results were published in the high-impact journals and do contribute to the state of the art. In fact, the quality of publications is extraordinary and well confirms the quality of the work. Developed theory of percolation may be used in a wide range of applications. Saeed is a first author of the paper "Optimal percolation on multiplex networks", which was published in the Nature Communications, a Nature Index high-impact journal. He is second author in another Nature indexed paper "Characterizing the analogy between hyperbolic embedding and community structure of complex networks" published in PRL, which demonstrates the recognition on the international level. I have positive opinion about the research contents of the thesis but I think that the text deserves some improvements (see issues below). To sum up, I think that the issues found do not decrease the scientific quality of the thesis and Saeed Osat deserves to be awarded with Skoltech PhD degree. The list of minor issues:

1. "list of publications": List of publications should contain full bibliographic references including pages. Currently it sometimes contains strange numeric codes instead of pages.

2. "sub-subsections numbers" Currently the text contains (very) many sub-subsections. I think that numbering them will slightly improve the experience of interacting with the text.

3. "Notations" Please, make sure that all the notations are introduced.

For example, one can guess that $<>$ is for averaging but anyway it should be introduced as this notation is not uniformly used across sciences.

I thank Prof. Panov for his report. Sub-subsections are now numbered and the notation of average is defined in its first appearance in the text.

List of changes in the revised thesis:

- **page 4.** In the 3rd column in the table the definition is updated.
- **page 5.** Definition of "Excess degree" and "Excess degree distribution" is added to the table.
- **page 15.** The whole section 1.1.11 is added in the revision, describing phase, phase transition and its types.
- **page 18.** The whole section 1.1.12 is added in the revision, describing the method of generating function.
- **Page 22.** Definition of susceptibility and its relevance to transition is added.
- **page 24.** Theoretical prediction for the phase transition of GMCC is added.
- **page 27.** Caption of the Figure 1.4 is updated.
- **page 28.** The analogy between ICM and BP is explained.
- **page 35.** P^{∞} for observability model is defined.