Summary of Habilitation Thesis

Bridging the Gap Between Computational Network Science and Computer Engineering

Alexandru Topîrceanu, PhD

Network Science is an interdisciplinary field by nature, with its roots embedded in mathematics, physics, statistics, computer science and information technology. Its applicability, however, stretches over to encompass the biological, pharmacological, social, economic, political sciences altogether. Today, Network Science research can be classified into the following fundamental topics: social network analysis (SNA), computational network analysis (CNA), biological network analysis, multilayer networks, dynamic network analysis, link (prediction) analysis, and centrality (influence) analysis.

The rapid rise of Network Science, over the last two decades, including many of today's modeling methodologies – involving large amounts of data and complex systems – are supported by Computer Engineering and Information Technology. As such, Information Technology offers approaches where computer algorithms, simulation tools and databases are used for the processing and understanding of biological, medical, pharmacological, or social data. Recent developments in personalized education and medicine are based on Big Data Analytics and Visualization, and Computer Science technologies such as Complex Network Analysis (CNA) and Machine Learning. Also, advancements in online social network technologies have enabled social physics and computational epidemics with global scale socio-economical impact.

The main motivation of the research presented in this thesis, spanning from 2011-2021, is to bridge the main field of Computers & Information Technology with the multidisciplinary field of Network Science through real-world, impactful applicability. This fusion results in the recently coined term of Computational Network Science, motivated by the fact that Computer Science and Engineering can change all other sciences through its data-driven approach. In support of this recent trend, we find new journals being dedicated to Computational Network Science: Nature Publishing Group has launched a new journal named *Nature Computational Science* in 2021 encouraging the development and use of computational procedures and mathematical models, including their utilization to solve complex problems across various scientific disciplines. Also, IEEE has launched the journal *IEEE Transactions on Computational Social Systems* addressing the modeling, analysis, simulation and understanding of social systems from a quantitative or computational perspective. Additionally, it is worth noting, that among important personalities, considered within the top 10 most powerful data scientists in the world, we find mostly computer engineers within academia (e.g., Stanford, MIT) and the industry (e.g., Google).

Given the high scientific impact and novelty of this field, we started research in the field of Network Science in Politehnica University Timisoara back in 2011. During my PhD studies (2012-2016) we published the first scientific results using a novel computational network approach, culminating with my PhD obtained in *Computers*. Thus, since 2016, we have laid the first bricks of a new School of Network Science within the Polytechnic context, and possibly the whole Romanian academic context.

Our research has extended along multiple tracks, such as social networks analysis (SNA), computational network analysis (CNA), network medicine. A first noteworthy contribution was to study the generation of realistic complex network topologies, their growth in time, followed by the opinion diffusion over large evolving networks. For this, we analyze state of the art topological models, inspired by the small world and scale free networks, aiming to reproduce empirically observed properties of real-world networks. To this end, we create a highly realistic social network model using a multi-variable genetic algorithm approach. Also, in contrast to the fundamental Degree preferential attachment principle, advocated by A-L. Barabasi, we further proposed the concept of *Betweenness preferential attachment* for better explaining the growth of social networks. Subsequently, we studied opinion spreading models using discrete and continuous opinion, including ones which include stubborn agents. Given the complex interplay of agent nodes, we chose computer simulation as a valid research methodology to evaluate and quantify these opinion spreading models over large social networks, because an analytical approach is not able to handle the high complexity of social network inter-connectivity. When applying the analytical power of computer modeling and simulation, we show that our novel *tolerance-based* model for opinion interaction and spreading generates realistic, reproducible patterns in social networks. While several existing models use "trust" parameters for agent interaction, we define "tolerance" as a time variable parameter dependent on the interaction patterns with neighboring nodes. Our inspiration for the evolution of tolerance derives from the idea that the dynamics towards tolerance and intolerance vary exponentially, meaning that an agent under constant influence becomes indoctrinated at an increased rate over time. Along this track, we also introduced a statistical tool for measuring the structural similarity (fidelity) between any two complex networks; modeled complex network antifragility under sustained attack ; and, introduced an original, reliable methodology for benchmarking node centrality measures in a competitive context. There are over 50 different node centralities used for the selection of spreader nodes in networks, and our methodology can reliably determine the more efficient ones given a specific topology.

Another important research track detailed in this thesis is the application of Network Medicine (i.e., Precision Medicine, Systems Medicine) providing computer-based solutions for medical and pharmacological challenges. On one hand, we worked for the past eight years in sleep research, in order to offer computational solutions for predicting the severity and development of Obstructive sleep apnea (OSA) and Chronic obstructive pulmonary disease (COPD). The results are formed around a patient phenotype model, built through a dual clustering technique, and a score usable by doctors in everyday patient monitoring. More recently, we extended the state of the art in OSA severity monitoring by presenting a differentiated phenotype model for each gender, as well as analyzing the causes of improved

CPAP treatment response. All these results are aimed at developing personalized treatment and precise diagnostics (like 4P Healthcare). On the other hand, we use the public database Drugbank to build a drug-drug interaction (DDI) network, where nodes represent drugs and links represent drug-drug interaction relationships between the drugs. More recently, we explored the potential of target based DDI, using the same dual clustering technique. These results help researchers estimate possible new interactions and repurposing alternatives for drugs, thus optimizing costly and time-consuming pharmacological studies.

Another undertaken track of interdisciplinary research has been that of adapting a complex network approach in the analysis of educational data. Specifically, we analyzed data of Romanian learners participating in MOOCs, and created a compatibility network of such students, based on their motivations, expectations, and perceived difficulties throughout the courses. By applying clustering techniques on the network, we defined specific student archetypes (i.e., corresponding to communities). Furthermore, we developed a fully original gamification platform for student motivation in class, and conducted an exam cheating study.

Our scientific and academic results are summarized by the management of 2 national research projects (financed by UEFISCDI), membership in an additional 2 international projects (financed by Linde and Horizon 2020), and 5 national projects (financed by UEFISCDI and ARUT), publication of 3 books, over 50 Web of Science indexed papers, out of which 16 journals (12 indexed in Q1/Q2), a cumulative impact factor over 45, a WoS h-index of 9 and 171 citations (330 citation in Google Scholar), review in diverse multidisciplinary and IEEE journals, organization of the 9 editions (2013-2021) of the SCMUPT student competition for mobile development, coordination as member in one PhD committee, and member of multiple PhD report committees, and coordination of over 90 Bachelor and Masters theses.

The currently available computing infrastructure, as published on the ERRIS platform, supports research on modeling and simulations, network science, big data, graph algorithms, and data mining, all of which are directly contributing research fields to our group. Furthermore, the infrastructures offered by the ACSA laboratory, the Department's Vision NextCloud platform, the University Virtual Campus, and the future available CloudPUTIng high performance computing platform will offer PhD students more than enough support for a diverse teaching and research career.

With our experience in coordinating PhD students over a broad range of scientific topics, we consider our ACSA (Advanced Computing Systems and Architectures) research group as an attractive opportunity for PhD programs. With the newest inclusion of Network Science in our research portfolio, we are able to offer a completely developed PhD program focusing on Network Science in the field of Computers and Information Technology.

Our future research will focus on two more recent research tracks (as of 2018), that of computational epidemics and political poll prediction. Given the COVID-19 pandemic, we were quickly motivated to collaborate with the University of Texas and focus our efforts on modeling and better understanding the impact of this outbreak. Ongoing research is under development on two tracks: that of understanding the impact of isolation strategies adopted in early 2020 and improving the underlying population model used for epidemic simulations.

Also, given the turmoil caused by local and global elections over the past years, we focused on improving the accuracy of pre-election polls using time series analysis and network science. We obtained encouraging results compared to existing state of the art methods, like Multilevel regression with poststratification.

In parallel, we aim to submit new project proposals, such as a TE (Tinere Echipe) and PED (Proiect Experimental Demonstrativ) project. Given the current ACSA members, we find both alternatives feasible. The TE can be targeted in the field of computational epidemics (a current impactful hot topic sub-field of computational network analysis). The PED may be targeted at the implementation of a sleep research related tool (for OSA). Also, the recently started Horizon 2020 project (2021-2025), in which several members of the ACSA and UMFT team are involved, opens endless possibilities for current master's students, as well as current and future PhD students. Given the large number of research centers involved in the project, most research publications will have a solid scientific impact, this opening opportunities of visibility for all participants.

All the mentioned scientific achievements illustrate the high potential of the thesis candidate as an independent researcher and his ability to manage research teams and future PhD students. Furthermore, he is actively collaborating with several research groups, with some notable results: Carnegie Mellon University (CMU) / University of Texas, University of Southern California (USC), and Central European University (research visit supported by the IMPRESS project). In the near future we propose establishing collaborations with the DSG group at Wroclaw University of Science and Technology, the Collide group at Universitaet Duisburg-Essen, and the DNDS department at CEU. These scientific collaborations should positively strengthen the international visibility, high originality of research, and relevancy of the work of ACSA group, as all mentioned teams work within the field of network science.

This thesis serves as a strong proof for the high impact research that can be achieved by employing Computer Science and Engineering in cross-disciplinary fields. As such, we intend to further narrow the gap between Computers and Network Science by further tackling challenging research topics from diverse fields of science, by participating at major conference venues in our field, by establishing long-lasting international collaboration, by creating project partnerships, by initiating new specialized undergraduate courses in our department, and by integrating new doctoral students in our ACSA research group.