MARTIN-LUTHER-UNIVERSITÄT
HALLE-WITTENBERG

Medizinische Fakultät
Institut für Medizinische Epidemiologie, Biometrie und Informatik (IMEBI)
Direktor: Prof. Dr. med. Rafael Mikolajczyk

22/06/2021

Review of the PhD thesis submitted by mgr inż. Tomasz Ożański in the fulfilment of the requirements for the degree of Doctor in the discipline "Information and Telecommunication" at Wroclaw University of Science and Technology

The review was assigned by the Scientific Council of the discipline "Information and Telecommunication" in connection with the proposed doctoral thesis prepared by mgr. inż. Tomasz Ożański with the title:

„Stochastic dynamics of systems with complex interactions
Dynamika stochastyczna układów ze złożonymi oddziaływaniami"

Haus-/Lieferanschrift:
Magdeburger Straße 8
06112 Halle (Saale)

Sekretariat:
Nadine Pietsch
Tel.: +49 (0)345 557-3570
Fax: +49 (0)345 557-3580
E-Mail:
imebmi@medizin.uni-halle.de

www.medizin.uni-halle.de/imebi

1. Topic of the thesis

The thesis focusses on the class of the models, which can be used for description of stochastic processes generally characterized as birth-and-death processes. In the doctoral thesis, the processes are illustrated by epidemic social contagious processes (spread of information in twitter network) and tumor development. The author postulates that combination of theoretical analysis with simulations leads to qualitatively and quantitatively new conclusions regarding the behavior of the system.

Maybe one could argue at this point that the title of the thesis could be more specific, for example referring to the class of processes studied in the thesis or to the connection between theoretical analyses and simulations, which are the central topic of the thesis.

The topic of studying stochastic dynamics is timely – while the current COVID pandemic raised interest in modelling of epidemic spread and lead to unprecedented involvement of mathematical modelers in political decision making regarding the choice of strategies to contain epidemics, also cancer remains a central challenge of health care systems, and mathematical cancer models are intensively researched worldwide. Overall, the in-silico models are increasingly used and became integral part of cutting-edge research in many areas of medicine and biology. Nevertheless, the development still has many areas in which only initial steps were made and further progress is needed. The thesis addresses important gaps and provides novel solutions.

2. Content of the thesis

Mr. Ożanski submitted his thesis in English language. The thesis is exactly 100 pages strong (event if single pages in between are left blank), and preceded by

a summary and explanation of main achievements, one page each, in Polish language. The thesis text consists of 11 chapters, with a bibliography containing 38 positions.

Chapter one is introduction to the thesis, and provides explanation of the considered class of birth-and-death processes and reference to the historical beginnings of the consideration of those in the Galton-Watson model. Furthermore, the aims and scope of the dissertation are presented and the main thesis is formulated. The introduction closes with a description of the structure of the text and explanation of the author`s achievements as contribution to the research.

Chapter two introduces the concept of the epidemics on random graphs. Starting with the Galton-Watson model, two more general, complex models (Erdös-Renyi and Bollobas-Janson-Riordan model) are introduced. The chapter closes with analysis regarding the giant component.

Chapter three addresses the question of time partitioning in social networks. The practical question is interesting and for example not sufficiently addressed in standard models of infections spread, possibly leading to false interpretations. While chapter two was just a description of the past developments, in chapter three starts the own contribution of the author. I have to admit that the transition between preexisting knowledge and new solutions is not well marked in the text. My preference can be linked to my relative unfamiliarity with the technical aspects, and possibly also specific writing traditions of the discipline, but from the more interdisciplinary perspective I would expect more explicit presentation what is already known and what is added knowledge. This is also very specific requirement of many journals. Apart from this technical aspect of presentation, the findings regarding communication time are relevant and interesting.

Chapter four extends the theoretical considerations of the Bollobas-Janson-Riordan network model. As in chapter three, it was difficult for me to identify the transitions between reporting what is already known and what is the new contribution of the author. This is explained in the introduction chapter, but it leaves the question open, if the extension regarding the direction of edges is a totally new idea or extended by the author. This can be easily solved by inserting short explanation in the text.

Chapter five addresses a more technical question of data preparation and storage. This is clearly an important aspect and I wanted to acknowledge positively that it was considered in the scope of the thesis. Here the introduction explicitly names the sections, which are new contributions by the author, and it can be inferred from the text. Having said that, I missed in this section the numerical results announced in 5.3.

Chapter six describes an application of the methods described in previous chapters to the epidemic spread of information in twitter networks. Figure 6 is missing color scheme, but an intuitive interpretation is still possible. In section 6.5, the central ideas of the thesis are addressed – here the stochastic simulations are compared with theoretical analyses and display a high agreement. It is also noted, that close to the critical value, the simulation model develops a mismatch to theoretical predictions, which have to be considered here as "true". At this point I am somewhat missing the consideration why a (less than perfect) simulation model is needed. I do not think that it is a fundamental problem, and

that it can be justified for example in the case of smaller samples or if individual characteristics are introduced, but it should be addressed in the context of the findings. The important finding is however that the extension of the BJR model towards the directional case substantially improved the agreement between theoretical considerations and simulation results. In this sense, the simulation is on the one side providing the support for a certain theoretical solution, but at the other side it can be explained why it is inferior at the margins of criticality.

Chapter seven builds a transition between the example of twitter data which is not spatially structured to the cancer example, which uses a n-dimensional space. As explained by the author, this chapter contains no own contribution and just reports the current state of knowledge. Helpful is the motivation section – explaining the biological mechanisms of tumor growth (section 7.3). It is well written and informative. In general, the section on tumor modelling is in my perception the strongest part of the work and also best organized in terms of explaining the gap in the knowledge and how it is addressed.

Chapter eight is introduced by the author as based on a jointly published manuscript, while indicating own contributions in the development of the model for verification of results, simulation algorithm and presentation of results in the manuscript. Before these parts are presented, the shape model for continuous space is introduced. The simulations conducted by the author extended in the area, which was not covered by the formal proof presented in the mentioned paper, and were performed to direct further research. Apparently, the findings stimulated further research, which resulted in another publication presenting a formal proof of the observation. This is a convincing example of research process in which the author was involved, and the fact that results became parts of publications is a positive quality indicator.

At this point, I became aware of the fact that while each chapter is well introduced, there is no true transition from preceding chapter. A help could be some sort of conclusion section in each chapter, particularly if new chapter covers a different topic. This would be even more apparent if chapters would immediately follow each other – not a new chapter starting at a new page. It can be clearly debated which is the best way of structuring such work, I can just report my personal impression that I missed transitions and closing sections in some of the chapters.

Chapter nine expands general remarks on tumor evolution from chapter seven. Here the author introduces a new model of tumor growth. Section 9.1 starts with "According to WHO, *Cancer is a generic term for a large group of diseases characterized by the growth of abnormal cells beyond their usual boundaries that can then invade adjoining parts of the body and/or spread to other organs.*" If the cursive means a direct citation – and it appears to be, then it should be enclosed in apostrophes, and a formal reference should be provided. This should be corrected when the thesis will be published. As for the figure 9.1 – the permission to use the figure should be provided in the figure caption, not just the reference. Figure 9.2. is probably own presentation developed by the author? In such case, figure caption should state (own presentation). The meaning of colors in figure 2 is not explained. While I like the description in Chapter 9.1, I have to say that in my discipline it would have to be much more intensively referenced. For example: "Even in a scope of the entire organism and multiple years, it seems impossible to gain all the required mutations just by chance. It is believed that the first mutations to be acquired give certain advantage to the

mutated cells over the cells that do not have the given mutation. The advantage will lead to increase in number of the cells with mutation due to Darwinian evolution. Another possible explanation might be that the first mutation acquired is the one responsible for increased mutation rate, eg. by weakening or disabling DNA repair mechanisms or allowing cells with damaged genome to proliferate, to allow for more mutations to take place in the future generations." – it is clearly not the authors own "possible explanation" and this should be made clear. I assume that the author follows here a textbook. This is acceptable, but should be explicitly stated and the text should be correctly paraphrased, i.e. making explicit, who provided the explanation etc..

Motivation in section 9.2 is very good again. Although also here sentences like "There already exist numerous simulation programs for tracking similar systems," should be amended by some references for exemplary chosen models. The remainder of chapter 9 is very systematic and informative.

Chapter ten provides the technical description of the algorithms developed by the author to implement the birth-and-death model in the time and space continuum. I would like to acknowledge specially, that the algorithm was published on GitHub repository. In this section, the high level of technical skills of the author is impressive to me.

Chapter eleven presents the findings from applying the model to tumor growth scenarios. In Figure 11.1 the text on the colored bar is too small. The findings are very interesting, for example regarding the surrounding layer of old clones. For figure 11.2, it is stated in the text that it is the cross-section presentation as in Fig 11.1 b, but this should be also mentioned in the figure caption.

The analysis of growth characteristics and phylogenetic trees is very interesting – I look very much forward to see it validated with real biological data. Coupled with experimental or clinical information, this model can provide very important insights into effects of various treatment regimens and their optimization. Findings presented in further figures provide very good insight into the abilities of the model.

Discussion section (11.4) is very useful (and is something I missed in section 6).

Chapter twelve provides concluding remarks, gives a well-balanced overview of the thesis and summary of findings.

Finally, I wanted to congratulate the author on the well written test, clear structure and organization of the thesis.

3. Evaluation of the achievements of the thesis

In the introduction, Mr. Ożanski pointed out several contributions to the field contained in the thesis. I am particularly impressed by the tumor evolution model. From my point of view, this part alone (chapter 9 and following) would be sufficient to declare an outstanding achievement. Clearly, this content requires introductory chapters 7 and 8. I also understand that the author was inclined to include in the thesis the smaller contributions described in the initial chapters and the social contagion process. Possibly, the tumor evolution model was more accessible to me, which resulted in the more appreciative view. As already commented, I had some difficulties in identifying the contribution of the

author in the initial chapters of the thesis, also for the social contagion model less explanations were provided and no discussion than for the tumor evolution model.

4. Suggestions and comments for future research

I think the tumor growth model has a strong potential for applications in further research, particularly as is seems to go beyond the current state of the art, therefore I would encourage the author to find experimental groups where the in-silico contribution will be welcomed.

The epidemic spread model received new importance due to COVID pandemics. The in the thesis presented application is maybe not very close to the real setting of pandemics and many models were developed for this particular purpose, so probably this should not be the focus. Nevertheless, the normal questions will return at some point, and the spread of information in social systems is relevant from many perspectives. As a potential theoretical extension, I could consider a system in which the process becomes multidimensional (i.e. different networks at the same time, with various overlap and various characteristics) interesting.

Conclusion

The thesis is scientifically rich, well conducted work. The candidate made several contributions to the field of stochastic models from informatics perspective, with the most extensive being the development and study of the continuous spatial model of tumor evolution.

Overall, Mr. Ożanski demonstrated in my view general theoretical knowledge in the field of informatics and his ability to perform independent scientific work. His thesis addresses relevant questions and presents substantial contribution to the research topic, and fulfills in my view the generally valid requirements for a doctoral thesis. I recommend the acceptance of the thesis by the responsible committee and propose the admission of the candidate to defense of the thesis.

Disclosure: I would like to explain that I sought to provide in the review the interdisciplinary perspective from medicine and epidemiology. Due to some modelling experience, I can also assess aspects of simulation modelling, but formal mathematical analysis is beyond my expertise. I am also not familiar with the customs of the Wroclaw University and topic specific – and had to extrapolate my reviewing experience from a different setting.

Sincerely

Prof. Dr. Rafael Mikolajczyk, MSc