



Poznań, June 26, 2022 r.

Mikołaj Morzy, Ph.D., D.Sc.

Mikolaj.Morzy@put.poznan.pl

Dissertation review

for the Council of the Faculty of Mathematics, Informatics and Mechanics, University of Warsaw

Title: Axiomatization of the Walk-Based Centrality Measures

Author: Tomasz Wąs

What scientific issue is considered in the paper (thesis of the paper) and has it been formulated clearly enough by the Author? What is the nature of the work (theoretical, experimental, other)?

The dissertation submitted for review is devoted to the axiomatization of a whole family of centrality measures based on feedback. In the dissertation the Author presents a set of axioms developed by him to characterize particular centrality measures and, based on the proposed axiomatic systems, proves very interesting properties of the considered centrality measures. The paper is mainly devoted to four popular centrality measures: PageRank, Seeley index, eigenvector centrality, and Katz centrality. The paper is theoretical in nature, a very large number of lemmas and theorems are included along with their proofs. The results presented in the paper are novel and very interesting. Both the methodology presented in the thesis and the theorems formulated are of far-reaching theoretical and practical significance. The Author presents many interesting and ingenious proofs related to the properties of the considered centrality measures and I am convinced that the presented approach may open a completely new research area in network science. The paper is original and interesting. Although difficult to read due to extensive mathematical formalism, the paper is prepared with great care and attention to the level of editing. I did not find any factual or methodological errors in the text of the dissertation.

The use of axiomatic systems in computer science has a long and fruitful history. Suffice it to mention the influence of the axioms of social influence and Arrow's theorem on the development of voting systems, or the success of relational database management systems often attributed to the axioms of concurrent data processing (the so-called transactional properties).



Today, when the field of machine learning, artificial intelligence or network science is dominated to such an extent by empirical approaches, the reviewed dissertation can be regarded as very innovative, undertaking the effort of painstaking theoretical considerations, with impressive practical consequences.

The problem of choosing a centrality measure to evaluate the importance of vertices in a network is a non-trivial problem. As the Author rightly points out, the choice of an appropriate measure must be dictated by both a good understanding of the semantics of network structures (in particular, the semantics of connections between vertices) and a thorough understanding of the properties and characteristics of individual centrality measures. Admittedly, the Author approaches the problem in a theoretical way, formulating a set of axioms and examining the properties of centrality measures in terms of satisfying these axioms, but it should be remembered that the conclusions of the analyses presented in the dissertation are of great practical importance. An example of such an elegant link between theory and practice is the identification of an axiom whose fulfillment guarantees the absence of incentives for individual vertices to manipulate the network structure in order to increase their centrality measure. No one needs to be convinced how important this property can be in relation to networks whose most "prestigious" vertices have a real impact on the functioning of the economy or society. The fact that, based on his analyses, the Author was able to propose a new measure of centrality that is much more resistant to manipulation than PageRank, I consider to be a very big achievement. In the light of the current state of knowledge in the area of research on network structures (the so-called network science), I am convinced that the topic discussed in the dissertation constitutes an important aim of scientific research, which fully meets the requirements for doctoral dissertations, and the theoretical results presented in the dissertation are original and important.

The dissertation is written in English and counts, together with the bibliography, 140 pages. It is accompanied by an abstract written in Polish. It consists of five chapters. The bibliography consists of 88 entries, 5 of which are co-authored by the PhD student and he is the first author. Chapter 1 is a short introduction. Chapter 2 defines the centrality measures considered in the paper, introduces auxiliary concepts used in proofs, and introduces all mathematical notations. Chapter 3 is devoted to the PageRank measure. The Author defines a set of six axioms and proves that they not only uniquely describe the PageRank measure, but that the PageRank measure is the only centrality metric that satisfies all six axioms. In Chapter 4, the Author introduced a new centrality measure, Random Walk Decay Centrality, which addresses the imperfections of the PageRank measure. For this purpose, the Author developed a different set of six axioms presenting more desirable properties and proved that the introduced measure satisfies them. An extremely interesting result presented in Chapter 5 is the relationship between



RWD and other centrality measures, in particular the decay centrality measure. The Author shows how, at the level of a set of axioms, one can switch between metrics by substituting a single axiom. Chapter 5 is an extension of the axiomatization attempts to the entire family of feedback-based centrality metrics. In this chapter, the Author introduces a third set of axioms and, in light of them, characterizes the metrics PageRank, Katz centrality, Seeley index, and eigenvector centrality. Chapters 3, 4, and 5 are extensive extensions of research papers presented at the International AAAI Conference on Artificial Intelligence and the International Joint Conference on Artificial Intelligence. Surprisingly, the dissertation does not have a chapter at the end, which would summarize and discuss the results collectively or suggest further directions for planned research.

Did the dissertation adequately review sources, including world literature, state of the art, and industry applications, demonstrating sufficient knowledge of the Author? Have the conclusions of the source review been formulated in a clear and convincing manner?

A subsection devoted to the presentation of the current state of knowledge is included at the very beginning of the dissertation. It is relatively short, but the Author skillfully avoids the danger of attempting to present the state of knowledge concerning measures of centrality (such a review could constitute a substantial monograph in itself). In the "Related work" section, the Author refers only to previous attempts to axiomatize centrality measures. Since measures based on distances and shortest paths are easier to analyze, most of the discussed work deals with such measures. In this review, the Author has a very ambitious goal of axiomatizing centrality measures based on influence, and the Author's review of the literature clearly shows that this research area has not been the subject of exhaustive work. In the review of previous research included in the paper, I only missed the reference to entropy-based centrality measures. Apart from that, I consider the selection of literature to be accurate, allowing the reader to properly place the dissertation against the background of previous research. The works cited are mostly contemporary, which proves the Author's orientation in the discussed subject matter.

Has the Author solved the questions posed, has he used the right method for doing so, and are the assumptions made justified?

As already mentioned, the work is purely theoretical in nature. The Author mentions possible practical consequences of the presented results, but the whole dissertation is devoted to proving a large set of lemmas and theorems supporting the introduced axiomatic system. I will not hide that detailed proofs of all the dozens of lemmas and theorems are complex and complicated, and I was not always (despite reading the text several times) able to fully follow the Author's thought.



Obviously, the consistency of the markings and the consistency of using a certain narrative template according to which the work should be read helped a lot in getting acquainted with the text. Undoubtedly, the Author used the right method (or rather set of methods) to prove the main theses of the dissertation. Slightly more difficult is the question about the validity of the assumptions made. By definition, axioms are not subject to discussion, but they should be chosen in such a way that all the theorems of a given theory follow from the chosen set of axioms. In the case of the reviewed work I have the impression that the Author uses the term "axioms" in the sense more similar to "postulate" or "properties". He defines some set of desirable properties and then analyzes particular measures of centrality in terms of satisfying these properties. From this perspective, of course, it is difficult to assess the validity of the axioms adopted and to state unequivocally whether the set introduced in the dissertation is the best possible one. Nevertheless, the justifications cited by the Author for the choice of particular axioms are convincing, and I am prepared to conclude that each of the three sets of axioms introduced is reasonable and practically useful.

What is the originality of the dissertation, what is the Author's independent and original achievement, what is the position of the dissertation in relation to the state of knowledge or level of technology represented by the world literature?

My overall assessment of the dissertation submitted for review is very high. I have no doubt that the considered problem is interesting and important, and the proposed solution is convincing and original. Importantly, the approach presented in the dissertation can be used to evaluate other families of centrality measures, and can also serve as a signpost for designing new measures that satisfy selected sets of axioms. It is difficult to clearly identify what constitutes the main contribution of the work. I would be inclined to believe that this contribution is the demonstration of the utility of axiomatization for such an important family of centrality measures as feedback-based measures. One can, of course, point to the new RWD metric as the main contribution of the dissertation, but it seems to me that the proposal of this metric is somehow a natural consequence of developing a set of axioms to describe the entire family of centrality measures. Reading the dissertation leaves no doubt that the Author's work is completely original and the contribution to the current state of knowledge is significant. The presented results are important and have great potential in the context of possible future research work.



Did the Author demonstrate the ability to correctly and convincingly present the results he/she obtained (brevity, clarity, editorial correctness of the dissertation)?

The editorial level of the dissertation is exemplary. Not only the work is completely free of spelling, grammatical and stylistic errors, but above all, the carefully planned structure of individual chapters is striking. The work is composed in a coherent and logical manner. I can see that a lot of work was put into organizing and making coherent the material from many conference publications. The result is an excellent text, which reads very well, despite its overload of lemmas, theorems and proofs. The mathematical designations used in the paper are used consistently and coherently. The quality of the English is exemplary, and the figures included in the dissertation are clear and chosen to visually illustrate portions of the text. If I were to compare the editorial level of this dissertation with dissertations I have reviewed over the years, I would place this dissertation in the top three. This makes the lack of a concluding chapter in the text all the more surprising.

What are the weaknesses of the dissertation and its major flaws?

Basically, the dissertation is devoid of any flaws. I did not find any substantive or methodological shortcomings. The editorial level of the work is exemplary, in 138 pages of text I did not come across a single error, typo, or even stylistically questionable construction. The consistency of the use of mathematical designations is astounding. Just the development of a set of mathematical designations in such a way that complex conditions considered in dozens of proofs can be expressed is awe-inspiring. Each chapter is carefully designed, with a clear introduction, development, and conclusion. The Author often helps the reader by guiding him through long mathematical deductions, hinting at the purpose of introducing the next lemma or at the conclusions to be drawn from proving a theorem. I am frankly amazed that it was possible to present such a complex argument so coherently and clearly.

I make a few comments below. They are not of a critical nature and do not in the least lower my very high evaluation of the reviewed text. These are the thoughts that came to my mind while reading the dissertation and they are at most an invitation to a discussion with the Author.

- On page 8, the Author introduces the terms "direct predecessor of v " and "direct successor of v ". I understand the need to distinguish vertices in a neighborhood based on edge direction, but I wonder if some adaptation of the term "neighbor" would not be more natural
- On page 19 the axiom "Node Deletion" is defined, the name seemed a bit confusing to me, wouldn't it be more accurate to call this axiom "Isolated Node Deletion"?



- I appreciate the tremendous work put into making mathematical notation more consistent. It would be all the more useful to have a summary table of the notations used in the dissertation, it would make the reading easier because many times during the analysis of the arguments I had to leaf through the dissertation looking for the definition of a given notation
- I think that resigning from the closing chapter was a mistake. An excellent text of the dissertation breaks off unexpectedly, without any summary, gathering once again the most important results. It leaves the reader in suspense, but not in a positive sense.

At this point I would like to formulate two more questions for the Author. First, in the light of the conducted research, does it seem possible to formulate a system of useful axioms for centrality indices that are a kind of aggregates? I mean indexes similar in construction to the Hirsch index, g-index or Impact Factor. Second, in the dissertation, a separate axiomatic system was introduced in each chapter. Of course, the individual axioms have been practically justified, but I am interested in whether the Author sees the possibility of defining more systematic relations between sets of axioms, in order to somewhat free himself from the subjective choice of individual axioms?

What is the usefulness of the dissertation to the engineering sciences?

The dissertation submitted for review solves an important theoretical problem and introduces an extremely useful tool for studying centrality measures in networks. The main contributions of the dissertation are:

- introducing a system of axioms uniquely characterizing the PageRank measure,
- proposing a new centrality measure RWD,
- introducing a consistent system of axioms to describe the four main centrality measures based on influence.

Each of these results would be interesting and innovative on its own, and their combination as a single dissertation is a major achievement. There is no doubt in my mind that the Author has excellent competence to conduct world-class research. The practical implications of the theses and claims presented in the dissertation can be very far-reaching. Partial results of the work have been presented at AAAI AI and IJCAI conferences, very prestigious industry conferences. I strongly believe that the ideas and solutions presented in the dissertation will have great resonance in the wider international scientific community.



POLITECHNIKA POZNAŃSKA

Poznan University of Technology

Doctoral dissertation of Tomasz Waś, "Axiomatization of the Walk-Based Centrality Measures" meets the conditions set out in Article 13(1) of the Act of 14 March 2003 on Scientific Degrees and Academic Title and Degrees and Title in Art (i.e. Journal of Laws of 2017, item 1789) in conjunction with Article 179(1) and (2) of the Act of 3 July 2018. Provisions introducing the Act - Law on higher education and science (Journal of Laws of 2018, item 1669, as amended). The reviewed work is an original and comprehensive solution to a scientific issue in the field of graph theory. **I evaluate the dissertation unequivocally positively** and request that it be admitted to public defense. Due to a very high, in my opinion, substantive level of the reviewed dissertation, supported by an excellent publication record of the PhD student, **I recommend that the Council of the Faculty of Mathematics, Informatics and Mechanics of the University of Warsaw distinguish the dissertation.**

A handwritten signature in blue ink, consisting of a stylized, cursive script.