



dr hab. inż. Dominik Strzałka, prof. PRz

Rzeszów, dn. 15.09.2024 r.

Politechnika Rzeszowska

Wydział Elektrotechniki i Informatyki

Zakład Systemów Złożonych

Tytuł rozprawy: Identification of hidden semantic relations in texts using relations' patterns

Autor rozprawy: mgr inż. Michał Gałusza

Promotor rozprawy: prof. dr hab. inż. Andrzej Walczak

Promotor pomocniczy:

Dziedzina: nauki inżynieryjno-techniczne

Dyscyplina: informatyka techniczna i telekomunikacja





Uwaga główna/General remark

Praca doktorska została (z wyjątkiem rozdziałów Podziękowania i Streszczenie) w całości napisana w języku angielskim. W związku z powyższym recenzja pracy została przygotowana w dwóch językach: polskim i angielskim.

The doctoral thesis (except for the Podziękowania and Streszczenie chapters) was written entirely in English. Therefore, the review of the thesis was prepared in two languages: Polish and English.





A synthetic presentation of the dissertation:

The main goal of the thesis

The main goal of the reviewed thesis is to present a solution for recognizing and modelling risk transmission that can be included in different documents that describe the operation or failure of a system. This is done in the context of whole documents and, in comparison to the so far existing approaches, based on risk analysis in sentences, is a significant extension in approaches for relationships identification. The flow of threats is considered as complex interactions between system elements; these interactions can be scattered between different documents.

Scope of the work

Generally, the scope of the work is focused on gradual decomposition of narrative descriptions and their presentation in several steps. The first one is based on the Semantic Frames Graph (SFG). Then the pattern threat propagation relationships are extracted to recognize propagation that is stored in the Intermediate Relationship Graph (IRG). They allow to do the final step to aggregate results in the form of Asset-Vulnerability-Hazard (A-V-H) graph. The dissertation presents a wide scope of theoretical considerations, practical examples, the original author proposal, and its evaluation by several examples.

Description of the work

The whole dissertation consists of six main chapters. In Chapter 1 which consists of 2 subsections, the author explains his motivation and outlines the challenges in risk analysis. The next chapter of the dissertation, with again 2 main subsections, gives a summary and limitations related to current methods from graph representation and text processing. Chapter 3 (with 4 main subsections) shows a wide discussion about NLP techniques to construct network representation of risk. Since the risk analysis is specific in many practical areas and there are no datasets that can be used as training sets, the author also focusses on existing bottlenecks. Chapters 4 (9 subsections) and 5 (5 subsections) are the most important ones. In Chapter 4 the author discusses a proposed solution with main concepts: Semantic Frame Graphs, Dialog Coherence, Intermediate Relationship Graphs. Chapter 5 shows the main results of the dissertation with a possible application to other areas. Chapter 6 shows the general summary.

It can be easily seen that the whole dissertation is divided into two major parts; Chapters 2 and 3 give theoretical background, while Chapters 4 and 5 present a practical solution and its evaluation.





The chapters layout is correct and gives impression to the reader that structure of dissertation is well-designed and well thought out.

Review about the presentation of general theoretical knowledge:

- the correctness of the originality of the thesis and the extent to which it has been demonstrated

In the dissertation there is no direct statement (sentence) of the main thesis. Author rather focusses on four research goals (page 13):

- ... answer if it is possible to construct a risk detection system without creating a dedicated LLM for Risk Analysis;
- the evaluation of available trained and language-specialized classifiers and construct the pipeline to identify risk relations without a dedicated training set;
- to provide a validation method;
- the final result shall produce an A-V-H graph as a network model of risk interaction.

Among these four research goals, the first one somehow represents the work thesis that is a kind of question: is it possible to construct a risk detection system without creating a dedicated LLM for Risk Analysis. The other three goals refer to confirmation of the stated question/thesis. On page 54 there is a problem statement (section 4.2) which solution is presented through the whole Chapter 4. The efficacy of the proposed solution is examined in Section 5. On page 78, in Section 5, the author wrote that it is possible to construct more comprehensive IRG and A-V-H graphs and it is done with relations to several real scenarios.

On pages 57-58 the author gives the proposed solution architecture and in fact for the first time stated the research questions. They are somehow recalled on page 76 where the author stated that solution proposed in Section 4 achieved all research goals. Little bit earlier, page 53, the author stated explicitly what is the main research problem: identifying risk interaction in the narrative. Furthermore, on page 76, the author, thanks to the time complexity analysis given in Section 4 was able to find and confirm that the proposed solution achieves the research goals.

From the formal, scientific, and practical point of view, this approach is acceptable and confirms the correctness of thesis originality and the extent to which it has been demonstrated.

- the analysis of sources

The author used references in the dissertation to 93 different sources. They are listed in the order as they appear in the text. Depending on the needs and context of considerations, they refer to quite wide period of time, but more than 50% were published during the last 10 years.





Taking into account the wide scope of the topic analyzed and the interdisciplinarity of risk analysis with current progress, especially in LLMs presented sources, is more than satisfactory. References are listed on pages 113 – 123 and for a small part of them there are missing some bibliographic details, especially the number of volumes, the number of pages: [1], [11] – [17], [26], [32-34], [38], [56], [57], [60] [62-66], [86-90].

- the position of the dissertation in relation to the state of knowledge and/or state of the art represented by world literature

The scope of work is related to very important practical problems that in general can be expressed as risk management. In the case of this special challenge, which is known for ages, humans were able to create different approaches, but today with the significant development of information processing based on computers and broadly understood artificial intelligence, it is expected that directly from narrative descriptions of different cases, situations, and problems, we will be able to analyze risk and create communicative solutions (for instance represented by graphs) integrating diverse perspectives. Obviously, the risk analysis can be a particular example of a more general problem that is expressed as identification of semantic relations in text via patterns recognition. In successive dissertation chapters, especially in Chapters 2 and 3, the author was able to show convincing analysis of related works in the context of ontology, graph representation of sentences (texts), network risk representation, entity recognitions, relationship extraction, textual entailment, semantic role labelling. This is a very strong part of the dissertation and not only gives an undoubtful convince that the author has very good topic understanding but more, that the rest of reviewed work could be very interesting and valuable. However, on pages 2-4 in the Introduction the author gives many important statements that could be supported by literature references. On page 2, the author shows a solution proposal, but the question appears: is this a novel, original author's proposal or it comes directly from the literature or, finally, is it a compilation of other approaches. If this is an original approach, it should be indicated in some way, but if not, it will be good to support this with literature references. On page 6 (1.2 Motivation), in Subsection 1.2.1 author recalls many supporting sentences and statements that justifies his interest in the thesis scope with relations to the real case scenarios and real technical problems but again they are not much supported by literature references.

The author was able to present his own solution that was also related and compared to other approaches; however, this could be done in a little bit more convincing way. It is mostly based on the DocRED dataset, which is used as a benchmark, and real-world examples. In my opinion, the results obtained could be further supported and confronted by results given in other scientific works and papers.





In summary, the comparison of the dissertation and its goals with the current state of knowledge and literature review gives strong conviction about valuable scientific work.

- opinion on the importance of the obtained results for a given scientific discipline

As written above, risk analysis and its transmission, directly from narrative documents that describe system operation or failures is not a simple challenge. The author was able to propose an approach where modelling risk propagation is done without dedicated relationship detection mechanisms. It is done with the use of Semantic Frame Graphs (SFG), then risk propagation is expressed by Intermediate Relationship Graphs (IRG) and finally aggregated by Asset-Vulnerability-Hazard (A-V-H) graph to do the network risk analysis. The proposed pipeline is new and can be used successfully in risk analysis. Because risk analysis is also a very important part of broadly understood computer engineering, in my opinion it is undoubtful that the result obtained by the author is very important for the information and communication technology discipline. This importance comes also from the fact that the author was able to achieve final results with the use of typical computer engineering tools and knowledge related to computer science.

In summary, this Ph.D. dissertation gives a valuable contribution to the discipline of information and communication technology.

- opinion about the author's ability to correctly and convincingly present obtained results

As was mentioned before, the entire dissertation is divided into two parts. The first one is related to the theoretical aspects and, in my opinion, is prepared very well. The author paid attention to many important issues and was able to convince the reader that despite some minor flaws mentioned above and below, in different parts of the review, his knowledge and experience related to the general scope are satisfactory. In the case of the second dissertation part, Chapter 4 is given in most of its parts without any doubtful issues.

However, in the case of Chapter 5 there are some things that should be further explained by the author. This is caused by at least two main reasons. The first one, a technical issue, is represented by the small size and low resolution of many figures, which should present confirmation of the obtained results. Figures are used for graphical presentations of risk flow analysis, and because the final result of document processing should be expressed as A-V-H graph, it is expected that these and other intermediate results will be given to the reader with required attention. The second one is caused, in my opinion, by the author's inattention – some of figures are not cited, some graph references are given wrong, figures are not described well and written sentences do not explain everything in detail; there is no efficient discussion. In the introduction to this chapter, the author wrote that previously presented theoretical foundations





will be empirically evaluated; even more, the author stated that: “This chapter is a critical juncture in my work,...”. More details about this will be given in the next parts of this review. Here, I would like to underline that these doubts can be explained and solved by the author if he will prepare answers for remarks given further in the review.

I would like to indicate that Chapter 6, where we have dissertation summary, can also be expanded because it does not present all conclusions that can be derived from the thesis. To be more precise: I am deeply convinced that the author put much more effort into this work preparation that can be concluded from this too short chapter. However, the important part of this chapter will also be indicated, the author's attitude towards continuing research and further work, but these statements do not fit the list of references where we have not received any author's own contribution.

The main disadvantages of the dissertation and its weaknesses.

Some of the dissertation disadvantages were presented above; here it is worth to note some other issues.

Remark 1: Page 24, we have Eq. (2.1) that is given without any justification or reference to external sources. I suppose that here the author relates the discussion to the term frequency–inverse document frequency (tf-idf) proposed by Karen S. Jones. There are several variants of the term frequency (tf) weight and also term inverse document frequency (idf). The tf-idf is calculated as the product of two statistics: term frequency (tf) and inverse document frequency (idf). There are various ways to determine the exact values of both statistics and, in the case of Eq. (2.1) it seems to be given as a weighting scheme, sometimes called count-idf. Please explain this.

Remark 2: Page 57, we have the function that expresses complexity for Triplet Validation Problem that is proven in appendix .1, but this is not stated in the text.

Remark 3: Page 61, we have Algorithm 1, but it is not directly stated who is the author. Please explain this.

Remark 4: On page 64, in Fig. 4.13 it is not clear how this figure corresponds to the example on page 63, especially to methapaths taking into account shapes (circles, diamonds, rectangles) and assumed notation, e.g. edges and directed edges.

Remark 5: Page 73, it is written: “Therefore, the complexity of a simple path between...”. I suppose that here the author wanted to say that this is a time complexity of single path generation, but this is not explained well in the dissertation. Please explain this.

Remark 6: Page 73. It is stated that the final [time] complexity is $O(N^3)$ but in my opinion if in the formula above the final result given as: $O(N) = N * (N - 1) *$





$(N - 2) \log(N - 2)$ and the sign $*$ refers to product, this calculation is wrong rather it should be $O(N^3 \log N)$. This should be explained.

Remark 7: Page 74, author wrote: "...in the IRG classification confirmed". I suppose that it is confirmed by LLM like in the case of $y_{(i)}^0$. Please explain this.

Remark 8: Page 80. The paragraph at the bottom of the page seems to be too enigmatic. First of all, the author wrote about dialogue-based strategies denoted as *isDialogRTE* and *isDialogRTE2*. Then refers to Fig. 5.3 and, further, it is stated that their dialog coherence score is low (Fig. 5.5) but for instance it is not directly related to any tables rows. So, there is a first question: How should this be interpreted? Next, we also have Fig. 5.6 that is not mentioned in the text, and it is hard to have a final imagination about the author's way of thinking. Please expand this.

Remark 9: Pages 82-84 show Figs. 5.8-5.10 and 5.12-5.13 that are given without any reference in the text and explanation. Some of them (it was mentioned below as a major flaw of the thesis) – Fig. 5.9, Fig. 5.11, Fig. 5.12, Fig. 5.13 have very low resolution or are very small to give the reader a general expression that everything what is presented here by the author is convincing. Please explain this.

Remark 10: Page 83. The description of Fig. 5.12 states: "New connections are highlighted.", but it can be seen that there are highlighted nodes but not connections. Please explain this.

Remark 11: Page 87. At the bottom, we have a confusion matrix (Table 5.1) but it is not commented on in the text. So what can be concluded particularly from this matrix in relation to the analyzed case?

Remark 12: Page 88, there is a Fig. 5.18 but again it is not mentioned in the text. What does it give in the context of an analyzed case? What else can be concluded from Fig. 5.19?

Remark 13: Page 98. The author wrote that: "The comparison of differences between the means for the models analyzed (Fig. 5.38) suggests using mistral-v.01. model". If we refer to the tables presented in Fig. 5.37 and Fig. 5.38 we have two rows with mistral-8x78-instruct-v0,1 and mistral-78-instruct-v0,2. I assume that the author is finally talking about mistral-8x78-instruct-v0,1 referring to the threshold from Fig. 5.37 equal to 0.273. But this is another example for which the author should pay more attention and be more precise especially in Chapter 5 where we have, as the author wrote, empirical evaluation and practical implementation. Please explain this.

Remark 14: Page 104. In Fig. 5.36 we have two histograms that at first sight look very similar, but their legend refers to flan_t5_large and flan_t5_xxl. I assume that they refer to results obtained from different LLMs, but how do these figures correlate to data given in Fig. 5.36 and Fig. 5.37?





Remark 15: Analysis and discussion given in Chapter 5 are unfortunately not well supported by included figures (also due to their very low resolution and small size) and many missing ones or given in a wrong way references. Chapter 5 provides the efficacy of the author's solution and should confirm its value. But the general impression after reading it is as follows: It is not convincing in many parts, or the author does not think over how to do this in the right way. On page 110 there is a summary presented in Section 5.6 – it is very short. It does not refer to the scope of very important results obtained in Chapter 5. It will be good to expand on it.

Remark 16: The final thesis summary (Chapter 6) also be longer with a better and more detailed emphasis on what was done in the work and how this refers to the current state of the art. If the reader refers to pages 13 and 78, where the author stated the main research goals and the dissertation thesis, only the first paragraph of Chapter 6 can be considered as confirmation and the direct author's statement that they were achieved and the thesis is confirmed. Please expand this.

Remark 17: How do the proposed solutions work in the case of the Polish language. In dissertation we have only examples related to English Language but, for instance, in Polish language nouns (and not only) are declined. How does this influence the performance and effectiveness of the proposed solutions pipeline?

Weaknesses: Some of the particular and general weak elements of the dissertation were indicated above, but here I would like to point out other remarks.

A major problem in this thesis is the small size and low resolution of many figures. For example:

- page 19, Fig. 2.2 (as a reproduction from [25]);
- page 25, Fig. 2.9 (as a reproduction from [29]);
- page 27, Fig. 12 (as a reproduction from [30]);
- page 74, Fig. 4.19 (as original figure by the author that provides important information on the validation and threshold calculation described in Sect. 4.9.1);
- page 79, Fig. 5.2 (as original, author's figure that confirms the existence of path in SFG decomposition and proves solution for example in Fig. 5.1);
- page 92, Fig. 5.21 (as original author's figure that gives the IRG graph eMARS example from Sect. 5.3);
- page 93, Fig. 5.22 (as original author's figure that shows the A-V-H graph for the eMARS example from Sect. 5.3);





- page 95, Fig. 5.24 (as original author figure showing interesting exception for eMARS example);

- page 95, Fig. 5.25 (as original author's figure that should confirm that the entailment score is above the assumed threshold);

- page 107, Fig. 5.41 (as original author's figure that gives the A-V-H graph for the financial scenario in Sect. 5.5.1);

- page 109, Fig. 5.45 and Fig. 5.48 (as original, the authors' figures that give the IRG graph and the A-V-H graph for medical scenario in Sect. 5.5.2).

Some of the other figures also deliver very important visualizations and justifications for different examples and statements, and sometimes it is very hard to notice this important element from figures because they are given with very small fonts.

Technical and language errors:

Page iv

konieczności stosowanie dedykowanego ... → konieczności stosowania dedykowanego

który pozwana na ... → który pozwala na

analizę spójnością dialogowej ... → analizę spójności dialogowej

nową metodą detekcji ... → nową metodę detekcji

Page 5

Chapter 2 (is given twice) → Chapter 3

Page 8

by the "engine" as the its flying capability ... → by the "engine" as its flying capability

Page 25

railway incident incident results... → railway incident results

Page 28

There are several examples of applications of such risk representation strategy. → There are several examples of such risk representation strategy applications.

Page 34

$P(y_i | x_{(i-k)}, \dots, x_{(i+l)}, y_{i-m}, \dots, y_{(i-1)}) \rightarrow P(y_i | x_{(i-k)}, \dots, x_{(i+l)}, y_{i-m}, \dots, y_{(i-1)})$

Page 66

$0 < f_{i,i+1}^c < 1 \rightarrow 0 < f_{i,i+1}^c < 1$ as it is stated on page 67 with explanations given under Eq. (4.1)





Page 66

(Example 2): and F_1 : $\rightarrow$ (Example 2): F_1

Page 67

Considerations related to computational complexity should be related to time complexity.

Page 72

modes will IRG modes $\rightarrow$ modes will be IRG nodes

Page 80

All specified methapaths perform the SFG graph (fig. 5.3) $\rightarrow$ All specified methapaths perform the SFG graph (Fig. 5.2) – wrong figure number

Page 97

... are as far as possible (Figs. 5.30, 5.32, 5.34). $\rightarrow$... are as far as possible (Figs. 5.30, 5.32, 5.35). Author refers to blue and orange curves that are visible on Figs. 5.30, 5.32, 5.35.

Page 97

... for each model (Figs. 5.31, 5.33, 5.34, 5.36). $\rightarrow$... for each model (Figs. 5.31, 5.33, 5.34, 5.36).

Pages 97-104

Figures could be renumbered to keep the order as they are referenced in subsection 5.4

Page 98

Big blank spaces before and after the equation.

Page 106

... „reasoning paths” (Figs. 5.42) 5.43). $\rightarrow$... „reasoning paths” (Figs. 5.42 and 5.43).

From a technical point of view, the thesis that was written in English looks very good. There are not many significant flaws, but this does not mean that they do not exist at all. The flaws noted above do not significantly influence the overall overview of the thesis.

Review about whether the doctoral dissertation demonstrate general theoretical knowledge

As was stated before, the author was able to support his work, especially in Chapters 2 and 3, by a very good theoretical background for the presented considerations. For proposed algorithms we have complexity analysis (one analysis requires short explanation – see Remark





6), author uses proper terminology, has a very good knowledge in theoretical background for scientific activities and demonstrated necessary amount of general theoretical knowledge together with many practical skills typical in computer science and engineering.

Review about whether the doctoral dissertation demonstrates the ability to independently conduct scientific work

In the dissertation the author presented two chapters (4 and 5) that are mainly focused on his independent and personal work. Chapter 4 presents a solution to a research problem: the main attention is paid to document-level relationship extraction. An original pipeline is constructed with several processing steps. This pipeline uses original author's solutions that from one hand are based on known approaches but also were expanded by original extensions relaxing the training set constraints. How this pipeline works is supported by experiments. Results are presented and briefly discussed. In the pipeline everything is done in multinomial-time complexity and this achieves the proposed research goals. The chapter content is presented in a very convincing manner and does not raise any doubts.

In the case of Chapter 5, as was mentioned in previous parts of the review, there are some doubts about the efficacy of the solution. It is not given as well as it would be expected (especially with relation to very positive expression from previous chapters), mainly due to some technical flaws that for instance refers to figures resolution, missing figures references and too short sentences that do not touch the core of expected details. This is a little bit disappointing, but finally it cannot be stated that significantly negatively influences the author's ability to independently conduct scientific work. Rather, this is related to some problems with appropriate presentation of obtained results. Despite this, in my humble opinion, I am convinced that any existing doubts can be explained by the author during the defense of the Ph.D. or by the answer to questions given in this review. I am convinced that the author has a full ability to conduct scientific work and in this particular case put a lot of effort in order to prepare the final results that confirm his original approach. More attention could be paid to a better presentation of the results.

Review about whether the doctoral dissertation is an original solution to a scientific problem, an original solution in the application of the results of one's own research in the economic or social sphere

As was written above and taking into account the content of the reviewed dissertation, the author was able to prepare and present an original solution to a scientific problem with an original and valuable practical solution that can be applied to real case scenarios and important real problems. The knowledge and practical skills presented are convincing, especially in the





context of risk analysis, which, as stated in the initial parts of this review, is a very important element of many human activities, starting from risk management, the flow of threats, and finishing in project management. Having in mind that many scenarios and real case problems are described in narrative, the proposed solution fits very well to broadly understood information processing that is a core of computer engineering, with particular attention paid to the processing of whole documents but not only single sentences.





Final conclusion

After the detailed analysis of the submitted doctoral dissertation prepared by MSc Eng Michał Gałusza, I would like to express that it makes a valuable contribution to the development of the information and communication technology discipline in the engineering and technology field (domain). It is an original solution to a scientific problem with original proposed practical solution of PhD candidate's scientific research, confirming that the doctoral dissertation presents general theoretical knowledge, very good substantive preparation, practical skills and the ability to independently conduct scientific research by a PhD student – Michał Gałusza. The subject matter is up-to-date and has a very high application potential. The reviewed doctoral dissertation meets the statutory requirements for candidates for the doctoral degree, specified in the Act on Academic Degrees and Titles - included in Art. 13 sec. 1 of the Act of March 14, 2003 on academic degrees and titles and degrees and titles in the field of art in connection with art. 179 of the Act of July 3, 2018. Provisions introducing the Act - Law on Higher Education and Science (Journal of Laws of 2018, item 1669).

Basing on this conclusion, I am applying to the IT and Telecommunications Discipline Council of the Military University of Technology in Warsaw for admission to its public defense.





Syntetyczna prezentacja rozprawy

Główny cel pracy

Głównym celem recenzowanej rozprawy jest przedstawienie rozwiązania do rozpoznawania i modelowania transmisji ryzyka, które może być zawarte w różnych dokumentach opisujących działanie lub awarię systemu. Odbywa się to w kontekście całych dokumentów i, w porównaniu do dotychczas istniejących podejść, opartych na analizie ryzyka w zdaniach, stanowi znaczące rozszerzenie podejść do identyfikacji relacji. Przepływ zagrożeń jest rozpatrywany jako złożone interakcje między elementami systemu; interakcje te mogą być rozproszone między różnymi dokumentami.

Zakres prac

Ogólnie rzecz biorąc, zakres pracy koncentruje się na stopniowym rozkładzie opisów narracyjnych i ich prezentacji w kilku krokach. Pierwszy z nich opiera się na grafie ramek semantycznych (Semantic Frames Graph, SFG). Następnie wyodrębniane są relacje propagacji zagrożenia wzorca, aby rozpoznać propagację przechowywaną w grafie pośrednich relacji semantycznych (Intermediate Relationship Graph, IRG). Pozwalają one na wykonanie ostatniego kroku, aby zagregować wyniki w formie grafu Asset-Vulnerability-Hazard (A-V-H). Rozprawa przedstawia szeroki zakres rozważań teoretycznych, przykłady praktyczne, oryginalną propozycję autora i jej ocenę na podstawie kilku przykładów.

Opis pracy

Cała rozprawa składa się z sześciu głównych rozdziałów. W rozdziale 1, który składa się z 2 podrozdziałów, autor wyjaśnia swoją motywację i zarysowuje wyzwania w analizie ryzyka. Następny rozdział rozprawy, również z 2 głównymi podrozdziałami, zawiera podsumowanie i ograniczenia związane z obecnymi metodami z zakresu reprezentacji grafów i przetwarzania tekstu. Rozdział 3 (z 4 głównymi podrozdziałami) przedstawia szeroką dyskusję na temat technik przetwarzania języka naturalnego w celu konstruowania reprezentacji sieciowej ryzyka. Ponieważ analiza ryzyka jest specyficzna w wielu praktycznych obszarach i nie ma zestawów danych, które można by wykorzystać jako zestawy treningowe, autor skupia się również na istniejących wąskich gardłach. Rozdziały 4 (9 podrozdziałów) i 5 (5 podrozdziałów) są najważniejsze. W rozdziale 4 autor omawia proponowane rozwiązanie z głównymi koncepcjami: semantyczne wykresy ramowe, spójność dialogu, pośrednie wykresy relacji. Rozdział 5 przedstawia główne wyniki rozprawy z możliwym zastosowaniem w innych obszarach. Rozdział 6 przedstawia ogólne podsumowanie.





Łatwo zauważyć, że cała rozprawa jest podzielona na dwie główne części; rozdziały 2 i 3 przedstawiają tło teoretyczne, podczas gdy rozdziały 4 i 5 przedstawiają praktyczne rozwiązanie i jego ocenę.

Układ rozdziałów jest poprawny i sprawia wrażenie u czytelnika, że struktura rozprawy jest dobrze zaprojektowana i przemyślana.

Przegląd dotyczący prezentacji ogólnej wiedzy teoretycznej

- poprawność oryginalności pracy i stopień jej wykazania

W rozprawie nie ma bezpośredniego sformułowania (zdania) głównej tezy. Autor skupia się raczej na czterech celach badawczych (strona 13):

- ... odpowiedź, czy możliwe jest zbudowanie systemu wykrywania ryzyka bez tworzenia dedykowanego LLM do analizy ryzyka;
- ocena dostępnych wyszkolonych i językowo wyspecjalizowanych klasyfikatorów oraz zbudowanie potoku do identyfikacji relacji ryzyka bez dedykowanego zestawu treningowego;
- dostarczenie metody walidacji;
- końcowy wynik ma wytworzyć graf A-V-H jako model sieciowy interakcji ryzyka.

Spośród tych czterech celów badawczych pierwszy w jakiś sposób reprezentuje tezę roboczą, która jest rodzajem pytania: czy możliwe jest zbudowanie systemu wykrywania ryzyka bez tworzenia dedykowanego LLM do analizy ryzyka. Pozostałe trzy cele odnoszą się do potwierdzenia postawionego pytania/tezy. Na stronie 54 znajduje się stwierdzenie problemu (sekcja 4.2), którego rozwiązanie jest przedstawione w całym rozdziale 4. Skuteczność proponowanego rozwiązania jest badana w sekcji 5. Na stronie 78, w sekcji 5, autor napisał, że możliwe jest skonstruowanie bardziej kompleksowych grafów IRG i A-V-H i robi to w odniesieniu do kilku rzeczywistych scenariuszy.

Na stronach 57-58 autor podaje proponowaną architekturę rozwiązania i faktycznie po raz pierwszy sformułował pytania badawcze. Są one w jakiś sposób przywoływane na stronie 76, gdzie autor stwierdził, że rozwiązanie zaproponowane w sekcji 4 osiągnęło wszystkie cele badawcze. Trochę wcześniej, na stronie 53, autor wyraźnie stwierdził, jaki jest główny problem badawczy: identyfikacja interakcji ryzyka w narracji. Ponadto na stronie 76 autor, dzięki analizie złożoności czasowej podanej w sekcji 4, był w stanie stwierdzić i potwierdzić, że proponowane rozwiązanie osiąga cele badawcze.

Z formalnego, naukowego i praktycznego punktu widzenia podejście to jest akceptowalne i potwierdza poprawność oryginalności tezy i stopień, w jakim została ona wykazana.





- analiza źródeł

Autor wykorzystał w rozprawie odwołania do 93 różnych źródeł. Zostały one wymienione w kolejności pojawiania się w tekście. W zależności od potrzeb i kontekstu rozważań odnoszą się one do dość szerokiego przedziału czasowego, ale ponad 50% zostało opublikowanych w ciągu ostatnich 10 lat. Biorąc pod uwagę szeroki zakres analizowanego tematu i interdyscyplinarność analizy ryzyka przy obecnym postępie, zwłaszcza w prezentowanych źródłach dotyczących LLM, jest to więcej niż zadowalające. Referencje zostały wymienione na stronach 113–123 i dla niewielkiej części z nich brakuje pewnych szczegółów bibliograficznych, zwłaszcza numeru tomów, numerów stron: [1], [11]–[17], [26], [32–34], [38], [56], [57], [60] [62–66], [86–90].

- pozycja rozprawy w stosunku do stanu wiedzy i/lub stanu sztuki reprezentowanej przez literaturę światową

Zakres pracy dotyczy bardzo ważnych problemów praktycznych, które generalnie można wyrazić jako zarządzanie ryzykiem. W przypadku tego szczególnego wyzwania, znanego od wieków, ludzie byli w stanie tworzyć różne podejścia, ale dziś, wraz ze znacznym rozwojem przetwarzania informacji opartego na komputerach i szeroko pojętej sztucznej inteligencji, oczekuje się, że bezpośrednio z opisów narracyjnych różnych przypadków, sytuacji i problemów będziemy w stanie analizować ryzyko i tworzyć rozwiązania komunikacyjne (na przykład reprezentowane przez grafy) integrujące różne podejścia. Oczywiście jest, że analiza ryzyka może być szczególnym przykładem bardziej ogólnego problemu, który wyraża się jako identyfikacja relacji semantycznych w tekście poprzez rozpoznawanie wzorców. W kolejnych rozdziałach rozprawy, szczególnie w rozdziałach 2 i 3, autor był w stanie przedstawić przekonującą analizę powiązanych prac w kontekście ontologii, reprezentacji graficznej zdań (tekstów), reprezentacji ryzyka sieciowego, rozpoznawania encji, ekstrakcji relacji, implikacji tekstowej, etykietowania ról semantycznych. Jest to bardzo dobrze wykonana część rozprawy i nie tylko daje niewątpliwe przekonanie, że autor ma bardzo dobre zrozumienie tematu, ale co więcej, że reszta recenzowanej pracy może być bardzo interesująca i wartościowa. Jednak na stronach 2–4 we Wprowadzeniu autor podaje wiele ważnych stwierdzeń, które można było poprzeć odniesieniami do literatury. Na stronie 2 autor przedstawia propozycję rozwiązania, ale pojawia się pytanie: czy jest to nowatorska, oryginalna propozycja autora, czy pochodzi bezpośrednio z literatury, czy wreszcie jest to kompilacja innych podejść. Jeśli jest to oryginalne podejście, należy to w jakiś sposób wskazać, ale jeśli nie, dobrze będzie poprzeć to odniesieniami do literatury. Na stronie 6 (1.2 Motivation), w podrozdziale 1.2.1 autor przywołuje wiele zdań i stwierdzeń pomocniczych, które uzasadniają jego zainteresowanie zakresem rozprawy w odniesieniu do rzeczywistych scenariuszy przypadków i rzeczywistych problemów technicznych, ale ponownie nie są one w dużym





stopniu poparte odniesieniami do literatury. Autor był w stanie przedstawić własne rozwiązanie, które było również powiązane i porównane z innymi podejściami; jednakże można to zrobić w nieco bardziej przekonujący sposób. Opiera się głównie na zbiorze danych DocRED, który jest używany jako punkt odniesienia, oraz na przykładach z prawdziwego świata. Moim zdaniem uzyskane wyniki można dodatkowo poprzeć i skonfrontować z wynikami podanymi w innych pracach naukowych i artykułach.

Podsumowując, porównanie rozprawy i jej celów z obecnym stanem wiedzy i przeglądem literatury daje silne przekonanie o wartościowej pracy naukowej.

- opinia o znaczeniu uzyskanych wyników dla danej dyscypliny naukowej

Jak napisano powyżej, analiza ryzyka i jego transmisja, bezpośrednio z dokumentów narracyjnych opisujących działanie systemu lub awarie nie jest prostym wyzwaniem. Autor był w stanie zaproponować podejście, w którym modelowanie propagacji ryzyka odbywa się bez dedykowanych mechanizmów wykrywania relacji. Odbywa się to przy użyciu grafów ramek semantycznych (SFG), a następnie propagacja ryzyka jest wyrażana za pomocą grafów relacji pośrednich (IRG) i ostatecznie agregowana za pomocą grafu zasób-podatność-zagrożenie (A-V-H) w celu wykonania analizy ryzyka sieciowego. Proponowany potok przetwarzania jest nowy i może być z powodzeniem stosowany w analizie ryzyka. Ponieważ analiza ryzyka jest również bardzo ważną częścią szeroko rozumianej inżynierii komputerowej, moim zdaniem nie ulega wątpliwości, że wynik uzyskany przez autora jest bardzo ważny dla dyscypliny informatyka techniczna i telekomunikacja. To znaczenie wynika również z faktu, że autor był w stanie osiągnąć wyniki końcowe przy użyciu typowych narzędzi inżynierii informatycznej i wiedzy związanej z informatyką.

Podsumowując, recenzowana rozprawa doktorska wnosi cenny wkład do dyscypliny technologii informatycznych i komunikacyjnych.

- opinia o umiejętności autora do poprawnego i przekonującego przedstawienia uzyskanych wyników

Jak wspomniano wcześniej, cała rozprawa podzielona jest na dwie części. Pierwsza dotyczy aspektów teoretycznych i moim zdaniem jest przygotowana bardzo dobrze. Autor zwrócił uwagę na wiele istotnych kwestii i potrafił przekonać czytelnika, że pomimo pewnych drobnych wad wymienionych powyżej i poniżej w różnych częściach recenzji, jego wiedza i doświadczenie związane z ogólnym zakresem są zadowalające. W przypadku drugiej części rozprawy, rozdział 4 jest podany w większości swoich akapitów bez żadnych poważniejszych wątpliwości.





Jednak w przypadku rozdziału 5 jest kilka rzeczy, które autor powinien nieco bardziej szczegółowo wyjaśnić. Jest to spowodowane co najmniej dwiema głównymi przyczynami. Pierwsza, kwestia techniczna, jest reprezentowana przez mały rozmiar i niską rozdzielczość wielu rysunków, które powinno stanowić potwierdzenie uzyskanych wyników. Rysunki służą do graficznej prezentacji analizy przepływu ryzyka, a ponieważ końcowy wynik przetwarzania dokumentów powinien być wyrażony jako graf A-V-H, jest oczekiwane, że te i inne pośrednie wyniki zostaną podane czytelnikowi z wymaganą uwagą. Drugim powodem jest, moim zdaniem, nieuwaga autora – część rysunków nie jest cytowana, niektóre odniesienia do wykresów są podane błędnie, rysunki nie są dobrze opisane, a napisane zdania nie wyjaśniają wszystkiego szczegółowo; nie ma efektywnej dyskusji. We wstępie do tego rozdziału autor napisał, że wcześniej przedstawione podstawy teoretyczne zostaną poddane empirycznej ocenie; co więcej, autor stwierdził, że: „Ten rozdział jest krytycznym momentem w mojej pracy,...”. Więcej szczegółów na ten temat będzie podanych w kolejnych częściach recenzji. Tutaj chciałbym podkreślić, że wątpliwości te mogą zostać wyjaśnione i rozwiązane przez autora, jeśli przygotuje odpowiedzi na uwagi podane w dalszej części recenzji.

Chciałbym też wskazać, że rozdział 6, w którym znajduje się streszczenie rozprawy, również można rozszerzyć, ponieważ nie przedstawia on wszystkich wniosków, jakie można wyprowadzić z rozprawy. Mówiąc dokładniej: jestem głęboko przekonany, że autor włożył znacznie więcej wysiłku w przygotowanie tej pracy, niż można wywnioskować z tego zbyt krótkiego rozdziału. Wskazano jednak również na istotną część tego rozdziału, czyli na postawę autora wobec kontynuowania badań i dalszej pracy, ale te stwierdzenia nie pasują do listy literatury, do której nie wniesiono żadnego własnego wkładu przez autora.

Główne wady rozprawy i jej słabe strony.

Niektóre z wad rozprawy zostały przedstawione powyżej; tutaj warto zwrócić uwagę na kilka innych kwestii.

Uwaga 1: Strona 24, mamy równanie (2.1), które jest podane bez żadnego uzasadnienia lub odniesienia do źródeł zewnętrznych. Przypuszczam, że tutaj autor odnosi dyskusję do ważenie częstością termów – odwrotna częstość w dokumentach (term frequency-inverse document frequency, tf-idf) zaproponowanej przez Karen S. Jones. Istnieje kilka wariantów ważenia częstotliwości termów (tf), a także odwrotnej częstotliwości dokumentów (idf). Tf-idf jest obliczany jako iloczyn dwóch statystyk: częstotliwości termów (tf) i odwrotnej częstotliwości dokumentów (idf). Istnieją różne sposoby określenia dokładnych wartości obu statystyk, a w przypadku równania (2.1) wydaje się, że jest ono podane jako schemat ważenia, czasami nazywany count-idf. Proszę to wyjaśnić.

Uwaga 2: Strona 57, mamy funkcję, która wyraża złożoność dla problemu walidacji potrójnej, co zostało udowodnione w dodatku 1, ale nie jest to stwierdzone w tekście.





Uwaga 3: Strona 61, mamy Algorytm 1, ale nie jest bezpośrednio powiedziane, kto jest autorem. Proszę to wyjaśnić.

Uwaga 4: Na stronie 64, na Rys. 4.13 nie jest jasne, jak ten rysunek odpowiada przykładowi na stronie 63, zwłaszcza w przypadku metaścieżek uwzględniających kształty (koła, romby, prostokąty) i założoną notację, np. krawędzie i krawędzie skierowane.

Uwaga 5: Strona 73, jest napisane: „Dlatego złożoność prostej ścieżki między...”. Przypuszczam, że tutaj autor chciał powiedzieć, że jest to złożoność czasowa generowania pojedynczej ścieżki, ale nie jest to dobrze wyjaśnione w rozprawie. Proszę to wyjaśnić.

Uwaga 6: Strona 73. Stwierdzono, że końcowa złożoność [czasowa] wynosi $O(N^3)$, ale moim zdaniem, jeśli w powyższym wzorze końcowy wynik podano jako: $O(N^3)$ a znak * odnosi się do iloczynu, to obliczenie jest błędne, raczej powinno wynosić $O(N^3 \log N)$. To powinno zostać wyjaśnione.

Uwaga 7: Strona 74, autor napisał: „...w klasyfikacji IRG potwierdzone”. Przypuszczam, że potwierdza to LLM, jak w przypadku $y_{(i)}^0$. Proszę to wyjaśnić.

Uwaga 8: Strona 80. Akapit na dole strony wydaje się być zbyt enigmatyczny. Przede wszystkim autor napisał o strategiach opartych na dialogu oznaczonych jako *isDialogRTE* i *isDialogRTE2*. Następnie odwołuje się do rys. 5.3 i dalej stwierdza, że ich wynik spójności dialogowej jest niski (rys. 5.5), ale nie jest to bezpośrednio powiązane z żadnymi wierszami tabeli. Tak więc pojawia się pierwsze pytanie: Jak należy to interpretować? Następnie mamy rys. 5.6, który nie jest wymieniony w tekście i trudno jest podążać z godnie ze sposobem myślenia autora. Proszę to rozwinąć.

Uwaga 9: Strony 82-84 pokazują rys. 5.8-5.10 i 5.12-5.13, które są podane bez żadnego odniesienia w tekście i wyjaśnienia. Niektóre z nich (zostało to wymienione poniżej jako główna wada tezy) – rys. 5.9, rys. 5.11, rys. 5.12, rys. 5.13 mają bardzo niską rozdzielczość lub są bardzo małe, aby dać czytelnikowi ogólne wrażenie, że wszystko, co jest przedstawione przez autora, jest przekonujące. Proszę to wyjaśnić.

Uwaga 10: Strona 83. Opis rys. 5.12 stwierdza: „Nowe połączenia są podświetlone.”, ale widać, że są podświetlone węzły, ale nie połączenia. Proszę to wyjaśnić.

Uwaga 11: Strona 87. Na dole mamy macierz pomyłek (Tabela 5.1), ale nie jest ona komentowana w tekście. Co zatem można konkretnie wywnioskować z tej macierzy w odniesieniu do analizowanego przypadku?

Uwaga 12: Strona 88, jest rys. 5.18, ale znowu nie jest on wspomniany w tekście. Co on daje w kontekście analizowanego przypadku? Co jeszcze można wywnioskować z rys. 5.19?

Uwaga 13: Strona 98. Autor napisał, że: „Porównanie różnic między średnimi dla analizowanych modeli (rys. 5.38) sugeruje użycie modelu mistral-v.01.”. Jeśli odwołamy się





do tabel przedstawionych na rys. 5.37 i rys. 5.38, mamy dwa wiersze z `mistral-8x78-instruct-v0,1` i `mistral-78-instruct-v0,2`. Zakładam, że autor ostatecznie mówi o `mistral-8x78-instruct-v0,1` odnosząc się do prognozy z rys. 5.37 równego 0,273. Ale to jest kolejny przykład niedokładności, na który autor powinien zwrócić większą uwagę i być bardziej precyzyjny, zwłaszcza w rozdziale 5, gdzie mamy, jak napisał autor, empiryczną ocenę i praktyczną implementację. Proszę to wyjaśnić.

Uwaga 14: Strona 104. Na rys. 5.36 mamy dwa histogramy, które na pierwszy rzut oka wyglądają bardzo podobnie, ale ich legenda odnosi się do `flan_t5_large` i `flan_t5_xxl`. Zakładam, że odnoszą się do wyników uzyskanych z różnych LLM, ale jak te liczby korelują z danymi podanymi na rys. 5.36 i rys. 5.37?

Uwaga 15: Analiza i dyskusja podane w rozdziale 5 nie są niestety dobrze poparte dołączonymi rysunkami (również z powodu ich bardzo niskiej rozdzielczości i małego rozmiaru), a do wielu z nich brakuje odniesień w tekście lub podano je w niewłaściwy sposób. Rozdział 5 przedstawia skuteczność rozwiązania autora i powinien potwierdzić jego wartość. Jednak ogólne wrażenie po jego przeczytaniu jest następujące: nie jest przekonujący w wielu częściach lub autor nie zastanowił się, jak to zrobić we właściwy sposób. Na stronie 110 znajduje się podsumowanie przedstawione w rozdziale 5.6 – jest bardzo krótkie. Nie odnosi się do zakresu bardzo ważnych wyników uzyskanych w rozdziale 5. Warto byłoby to rozwinąć.

Uwaga 17: Jak proponowane rozwiązania działają w przypadku języka polskiego. W rozprawie mamy tylko przykłady związane z językiem angielskim, ale na przykład w języku polskim rzeczowniki (i nie tylko) są odmieniane. Jak to wpływa na wydajność i skuteczność proponowanego rozwiązania?

Niektóre ze szczególnych i ogólnych słabych elementów rozprawy zostały już wskazane powyżej, ale tutaj chciałbym zwrócić uwagę na inne kwestie.

Głównym problemem w tej rozprawie jest mały rozmiar i niska rozdzielczość wielu rysunków. Na przykład:

- strona 19, rys. 2.2 (jako reprodukcja z [25]);
- strona 25, rys. 2.9 (jako reprodukcja z [29]);
- strona 27, rys. 12 (jako reprodukcja z [30]);
- strona 74, rys. 4.19 (jako oryginalny rysunek autora, który dostarcza ważnych informacji na temat walidacji i obliczania prognoz opisanych w sekcji 4.9.1);
- strona 79, rys. 5.2 (jako oryginalny rysunek autora, który potwierdza istnienie ścieżki w rozkładzie SFG i udowadnia rozwiązanie na przykład na rys. 5.1);





- strona 92, rys. 5.21 (jako oryginalny rysunek autora, który podaje przykład wykresu IRG eMARS z sekcji 5.3);

- strona 93, rys. 5.22 (jako oryginalny rysunek autora, który pokazuje wykres A-V-H dla przykładu eMARS z sekcji 5.3);

- strona 95, rys. 5.24 (jako oryginalny rysunek autora, który pokazuje interesujący wyjątek dla przykładu eMARS);

- strona 95, rys. 5.25 (jako oryginalny rysunek autora, który powinien potwierdzić, że wynik wnioskowania jest powyżej założonego progu);

- strona 107, rys. 5.41 (jako oryginalny rysunek autora, który pokazuje wykres A-V-H dla scenariusza finansowego w sekcji 5.5.1);

- strona 109, rys. 5.45 i rys. 5.48 (jako oryginalne rysunki autora, które pokazują wykres IRG i wykres A-V-H dla scenariusza medycznego w sekcji 5.5.2).

Niektóre z pozostałych rysunków dostarczają również bardzo ważnych wizualizacji i uzasadnień dla różnych przykładów i stwierdzeń, a czasami bardzo trudno jest zauważyć ten ważny element na rysunkach, ponieważ są one podane bardzo małą czcionką.

Błędy techniczne i językowe:

Strona iv

konieczności stosowanie dedykowanego ... → konieczności stosowania dedykowanego

który pozwana na ... → który pozwala na

analizę spójnością dialogowej ... → analizę spójności dialogowej

nową metodą detekcji ... → nową metodę detekcji

Strona 5

Chapter 2 (podano 2 razy) → Chapter 3

Strona 8

by the "engine" as the its flying capability ... → by the "engine" as its flying capability

Strona 25

railway incident incident results... → railway incident results

Strona 28

There are several examples of applications of such risk representation strategy. → There are several examples of such risk representation strategy applications.

Strona 34





$$P(y_i | x_{(i-k)}, \dots, x_{(i+l)}, y_{i-m}, \dots, y_{(i-1)}) \rightarrow P(y_i | x_{(i-k)}, \dots, x_{(i+l)}, y_{i-m}, \dots, y_{(i-1)})$$

Strona 66

$0 < f_{i,i+1}^c < 1 \rightarrow 0 < f_{i,i+1}^c < 1$ jak podano na stronie 67, z wyjaśnieniami pod równaniem (4.1).

Strona 66

(Example 2): and F_1 : $\rightarrow$ (Example 2): F_1

Strona 67

Rozważania dotyczące złożoności obliczeniowej powinny być odnoszone do złożoności czasowej.

Strona 72

modes will IRG modes $\rightarrow$ modes will be IRG nodes

Strona 80

All specified methapaths perform the SFG graph (fig. 5.3) $\rightarrow$ All specified methapaths perform the SFG graph (Fig. 5.2) – zły numer rysunku

Strona 97

... are as far as possible (Figs. 5.30, 5.32, 5.34). $\rightarrow$... are as far as possible (Figs. 5.30, 5.32, 5.35). Autor odnosi się do niebieskich i pomarańczowych krzywych widocznych na rys. 5.30, 5.32, 5.35.

Strona 97

... for each model (Figs. 5.31, 5.33, 5.34, 5.36). $\rightarrow$... for each model (Figs. 5.31, 5.33, 5.34, 5.36).

Strony 97-104

Rysunki można ponumerować ponownie, aby zachować kolejność, w jakiej są przywoływane w podsekcji 5.4.

Strona 98

Duże puste miejsca przed i po równaniu.

Strona 106

... „reasoning paths” (Figs. 5.42) 5.43). $\rightarrow$... „reasoning paths” (Figs. 5.42 and 5.43).

Z technicznego punktu widzenia praca napisana w języku angielskim wygląda bardzo dobrze. Nie ma wielu istotnych wad, ale nie oznacza to, że ich w ogóle nie występują. Wady wymienione powyżej nie wpływają znacząco na ogólny obraz pracy.





Opinia, czy rozprawa doktorska wykazuje ogólną wiedzę teoretyczną

Jak wspomniano wcześniej, autor był w stanie bardzo dobrze poprzeć swoją pracę, zwłaszcza w rozdziałach 2 i 3, w zakresie przedstawionych rozważań od strony teoretycznej. Dla proponowanych algorytmów mamy analizę złożoności (jedna analiza wymaga krótkiego wyjaśnienia – patrz Uwaga 6), autor używa właściwej terminologii, ma bardzo dobrą wiedzę w zakresie aspektów teoretycznych dla swojej aktywności naukowej i wykazał się niezbędnym zakresem ogólnej wiedzy teoretycznej wraz z wieloma praktycznymi umiejętnościami typowymi dla informatyki.

Opinia o tym, czy rozprawa doktorska wykazuje zdolność do samodzielnego prowadzenia pracy naukowej

W rozprawie autor przedstawił dwa rozdziały (4 i 5), które skupiają się głównie na jego osobistym wkładzie w rozprawę. Rozdział 4 przedstawia rozwiązanie problemu badawczego: główną uwagę poświęcono ekstrakcji relacji na poziomie dokumentu. Oryginalny potok przetwarzania został skonstruowany z kilkoma pośrednimi krokami przetwarzania. Ten potok używa oryginalnego rozwiązania autora, które z jednej strony opiera się na znanych podejściach, ale zostało również rozszerzone o autorskie koncepcje, które rozluźniają ograniczenia dotyczące zestawu treningowego. Sposób działania tego potoku jest poparty eksperymentami. Wyniki są przedstawione i krótko omówione. W potoku przetwarzanie odbywa się w wielomianowej złożoności czasowej, co oznacza osiągnięcie zaproponowanych celów badawczych. Treść rozdziału jest przedstawiona w bardzo przekonujący sposób i nie budzi żadnych wątpliwości.

W przypadku rozdziału 5, jak wspomniano w poprzednich częściach recenzji, istnieją pewne wątpliwości na temat skuteczności prezentacji rozwiązania. Nie jest ono podane tak dobrze, jak można by się spodziewać (zwłaszcza w odniesieniu do bardzo pozytywnych wyrażań z poprzednich rozdziałów), głównie z powodu pewnych wad technicznych, które na przykład odnoszą się do rozdzielczości rysunków, brakujących odniesień do rysunków i zbyt krótkich zdań, które nie dotykają sedna oczekiwanych szczegółów. Jest to trochę rozczarowujące, ale ostatecznie nie można stwierdzić, że ma to znaczący negatywny wpływ na zdolność autora do samodzielnego prowadzenia pracy naukowej. Raczej jest to związane z pewnymi problemami z odpowiednią prezentacją uzyskanych wyników. Mimo to, moim skromnym zdaniem, jestem przekonany, że wszelkie wątpliwości mogą zostać wyjaśnione przez autora podczas obrony doktoratu lub poprzez odpowiedź na uwagi podane w tej recenzji. Jestem przekonany, że autor ma pełną zdolność do prowadzenia pracy naukowej i w tym konkretnym przypadku włożył wiele wysiłku w przygotowanie końcowych wyników, które





potwierdzają jego oryginalne podejście. Więcej uwagi można by poświęcić lepszej prezentacji wyników.

Opinia o tym, czy rozprawa doktorska jest oryginalnym rozwiązaniem problemu naukowego, oryginalnym rozwiązaniem w zastosowaniu wyników własnych badań w sferze gospodarczej lub społecznej

Jak napisano powyżej i biorąc pod uwagę treść recenzowanej rozprawy, autor był w stanie przygotować i przedstawić oryginalne rozwiązanie problemu naukowego z oryginalnym i cennym rozwiązaniem praktycznym, które można zastosować do rzeczywistych scenariuszy przypadków i ważnych rzeczywistych problemów. Przedstawiona wiedza i umiejętności praktyczne są przekonujące, zwłaszcza w kontekście analizy ryzyka, która, jak stwierdzono w początkowych częściach tej recenzji, jest bardzo ważnym elementem wielu działań człowieka, począwszy od zarządzania ryzykiem, przepływu zagrożeń, a skończywszy na zarządzaniu projektami. Mając na uwadze, że wiele scenariuszy i rzeczywistych problemów przypadków jest opisanych w narracji, proponowane rozwiązanie bardzo dobrze wpisuje się w szeroko pojęte przetwarzanie informacji, które jest sensem informatyki, ze szczególnym uwzględnieniem przetwarzania całych dokumentów, a nie tylko pojedynczych zdań.





Wniosek końcowy

Po analizie przedłożonej rozprawy doktorskiej Pana mgra Michała Gałuszy stwierdzam, że wnosi ona wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja. Stanowi ona oryginalne rozwiązanie problemu naukowego, potwierdzając wiedzę, bardzo dobre przygotowanie merytoryczne, umiejętności praktyczne oraz umiejętność samodzielnego prowadzenia badań naukowych przez Doktoranta. Podjęta tematyka jest aktualna i ma bardzo duży potencjał aplikacyjny. Recenzowana rozprawa doktorska spełnia ustawowe wymagania stawiane kandydatom do stopnia doktora, określone ustawą o stopniach i tytułach naukowych – uwzględnione w art. 13 ust. 1 Ustawy z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki w związku z art. 179 Ustawy z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę - Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r., poz. 1669).

Na tej podstawie wnioskuję do Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Wojskowej Akademii Technicznej w Warszawie o dopuszczenie jej do publicznej obrony.

