Dr hab. inż. Jordi Mongay Batalla
Politechnika Warszawska
Jordi.mongay.batalla@pw.edu.pl

05 lutego 2024r.

Reviewer's opinion

on Ph.D. dissertation authored by

Bartosz Musznicki

entitled:

Multicast communication in wireless sensor networks with the use of uncontrolled mobile relay node

1. Problem and its impact

The dissertation presented by the PhD candidate Bartosz Musznicki has 154 pages and is fully written in English, except for the abstract translated into Polish.

The theme of the dissertation is multicast communications in urban wireless networks. Such networks are designed for measurement purposes and are characterised by heterogeneous structures comprising devices with diverse connectivity, computational, and functional capabilities. They typically include both stationary and mobile nodes, resulting in distributed and fragmented structures (due to the impossibility of predicting the physical localisation of the nodes).

In this situation, the response is to add packet forwarding functionality to the mobile nodes, so that the transmission between mobile nodes may increase communication reliability. This is acceptable as long as we are not dealing with mission-critical applications. The concept of delay-tolerant networks has been created to support this approach in wireless networks (including, and perhaps especially, urban networks). In delay-tolerant networks, the delay or latency of the packets is not an issue. These networks may include all IoT (Internet of Things) applications, which are the main applications in urban wireless networks. In delay-tolerant networks, mobile devices serve additional functions beyond their primary roles, often facilitating communication between disjoint areas. Packets are stored in relay nodes for future forwarding, with mobile relays facilitating communication between stationary subareas of the network.

The author proposes several multicast algorithms that differ in the knowledge that the nodes have about the network topology or the network structure. This can be without any knowledge, with local or global knowledge, and with current or predicted knowledge.

In the algorithm that does not assume any knowledge of the network situation, the candidate proposes that the nodes transmit each stored message during every algorithm iteration (time-slot). Subsequently, the message storage is refreshed, and the time-to-live (TTL) parameter of the stored message is decreased. Once the TTL reaches 0, the message is deleted from the storage. In this case, it does not seem possible to achieve much better performance (without any information on the network structure). In the case of the algorithm with local information, the main limitation is the mobility of the nodes. In conjunction with the dynamic changes in network structure resulting from the movement of mobile nodes, the locally available real-time knowledge about the current node's neighbours facilitates the multicast message's ability to prevent becoming trapped in a local optimum without progressing

toward desired destinations. The creation process of outgoing messages will commence only when there are currently neighbouring nodes present. In the case of global information, the node initiating the message has access not only to the current topology of the entire network but also to its anticipated future structure, enabling efficient message forwarding to desired destinations. This implies possessing global awareness of the network. Consequently, optimal multicast relays can be identified, as outlined by the author, for instance, using the shortest path algorithm, and then encoded in the message as the set of next hops known globally within the multicast tree.

One of the main parts of the dissertation is analysing real data provided mainly by cities. New public datasets are available, often in real-time. The author has used open data from public transportation and urban infrastructure in four Polish cities to develop a graph modelling method for representing time-changing network connectivity structures. This modelling approach forms the basis for architectural design and research environments, which were employed to study proposed multicast algorithms. The results demonstrate that uncontrolled mobile relays can be utilised to construct multicast structures that span both space and time, with their characteristics influenced by city topology, destination node distribution, and mobile relay routes. The efficacy and efficiency of these algorithms are contingent upon factors such as node radio range, maximum message forwarding duration, and the availability of network structure knowledge.

The problem posed by the Ph.D. candidate is important for networking due to two main reasons: on the one hand, it is unclear which wireless technology may be used in urban networks, so some kind of overlay solution is required, and, on the other hand, the necessity of more processing power in wireless ad-hoc networks has put emphasis on the necessity of the fog computing, which includes the transmission of a large amount of data.
In general, the developed algorithms are both practical and straightforward to implement. This means that the practical implementation of the proposed algorithms is high.
The way how the candidate has proposed and demonstrated the solutions follows a scientific method. However, there are some flaws in the work, as described in the next sections.

2. Contribution

The main contributions provided by the Ph.D. candidate are structured along the dissertation as follows (excluding the Introduction, which does not include any remarkable contribution and is only a description of the sections):

- Chapter 2 provides an overview of wireless sensor networks currently deployed in urban environments. This chapter traces the evolution of sensor networks and related concepts, examines the types and characteristics of existing sensor networks in urban settings, addresses key routing challenges in urban sensor networks, covering opportunistic routing, data aggregation, offloading, and topology control and modelling, and discusses pertinent observations;
- Chapter 3 analyses various node deployment schemes encountered in urban environments, illustrated with real-world examples. The section explores random node deployment, discusses deterministic node distribution, and outlines practical implications.
- Chapter 4 introduces an innovative method for modelling real-life urban sensor networks based on open data. In this section, the author examines the characteristics of urban node location data sources, introduces the architecture for data gathering, processing, and network modelling, presents novel network modelling algorithms with flow diagrams and pseudocodes, demonstrates the feasibility of

the introduced network modelling methodology, and summarizes the implications of preliminary verification research.

• Chapter 5 presents an urban delay-tolerant multicast framework using uncontrolled mobile relays. The author outlines key design considerations, introduces the new delay-tolerant multicast framework, comprising multicast message-oriented algorithms and procedures designed for opportunistic heterogeneous urban sensor networks with uncontrolled mobile relays, and defines and discusses the designed delay-tolerant multicast router, outlining key observations.

• Chapter 6 presents a simulation study of the introduced models and algorithms. It covers the research environment, simulation scope, architecture, and analysis methodology. This section discusses space connectivity issues, analyses space-time connectivity aspects, examines the results of space-time multicast analysis and highlights observations from the simulation study.

• Chapter 7 concludes with a summary, lists the main contributions, and suggests directions for further research.

I think that the contributions provided by the Ph.D. candidate go beyond the State of the Art. The most interesting parts of the dissertation are the analysis of massive open data and the proposed algorithms.

The candidate has analysed recent advancements and potential applications of publicly accessible real-world diverse data sources, serving as the foundation for an innovative approach to urban Wireless Sensor Network research. His approach relied on authentic node data and locations, departing from synthetic models or limited homogeneous historical data. The presented analysis opens avenues for modelling realistic, dynamic graphs representing spatial and temporal changes, which can be visualised, stored, and analysed to enhance various aspects of network design and optimisation.

The delay-tolerant multicast algorithms presented by the candidate for heterogeneous urban network structures prioritise practicality. They are designed to be protocol-independent, modular, and presented in a format conducive to understanding and implementation. While considering high-level aspects such as network utilisation, delivery ratio, computing power, storage, and wireless medium usage, intricate details are set aside to focus primarily on routing concerns. These methods are introduced as a family of interconnected modular components and procedures. Their distinguishing features lie in the utilisation of no, local, or global topology knowledge available to the nodes, as well as in the definition of multicast destinations, which can be single nodes, specific geographical regions, or selected node classes.

Sometimes, it is unclear if all the algorithms are a contribution from the candidate. For example, is the "*network devices data to slots of space nodes (NDD-SSN)*" an authored contribution?

The main lacks in the dissertation are, in my opinion, the following:

- There is no mathematical modelling or deeper analysis of the impact of the mechanisms;
- The proposed algorithms have not been deployed in a real network. This is quite understandable since urban wireless networks are not normally available beyond the use of the data that they generate;
- There are several practical issues when implementing the multicast algorithms that have not been considered by the author (see the next section), which limits the solution's impact on real network deployment.

However, I do not consider the aforementioned as real flaws but more as incompleteness of the work. In general, I do not have doubts that the work done by the Ph.D. candidate is a step forward in the analysis of urban wireless networks and, concretely, in the delivery of packets in urban networks when the delay is not a priority.

3. Correctness

The dissertation presents very interesting research on several points.
The author has done a great job analysing the data provided in open access. The data are from real measurements, which gives even more value to the presented results. The use of Machine Learning algorithms makes this analysis feasible, and the author has done hard work by handling and analysing such an amount of data.
The analysis of the algorithms based on the prediction of localisation is also an exciting way to provide more information for routing tasks.

There are, however, some flaws in the work, mainly due to the lack of practical implementation.
The author proposes in algorithms NKOM and LKOM that "each message includes the list of nodes visited by this message and current TTL value". The problem arises when we consider jumbo frames in the network. The author should discuss this case, in my opinion, since it can be a real barrier to the practical deployment of the algorithm in real networks. Similarly, it is unclear how such messages will be protected against integrity attacks. It seems quite logical to think of an attack that simply eliminates the list of visited nodes, and the message will run around the network. I understand that security is not the field of this dissertation but this point is an obvious problem of such an algorithm. Solutions such as JSON Web Signature try to solve those problems.
The visited nodes are also included in the GKOM algorithm. Still, in this algorithm, they are introduced during the transmission from the source (a priori knowledge), so the problem does not exist in this case.

It is highly appreciated that the author provides, when possible, average values or trending values. A minor flaw is that sometimes the confidence intervals are shown, but it is not clear at what confidence level the intervals were calculated (e.g., Figure 3.1).

I found also other small issues in the work, mainly related to the formatting and organisation of the dissertation. Many of them are due to the high quantity of data handled and analysed by the author in the text.
In general, the organisation of the dissertation is quite abnormal and difficult to follow sometimes. The analysis of the data is done in Section 4. Still, also in this section, there are many references to the algorithms that will be provided and described only in Section 5. Section 5 presents the algorithms, but the simulation analysis is done in Section 6, and this simulation analysis separates the data by cities more than by algorithms. This makes the reading quite difficult. A reader may ask why it is so important to show the differences between cities, more even than the differences between the algorithms (the differences in the algorithms are also there, but much more difficult to see than between cities). In other words, the candidate presents the conclusions of the results, taking as the reference the different parameters of the data in the different cities (more mobile relays, more node degrees, etc.), and this makes the impact of the parameters in the functioning of the algorithms is unclear many times.
A second problem with the organisation of the work is that the author has provided much code directly into the work. This really shows the intensive work that he has done; however, sometimes, one would expect more explanation of the code instead of the copy-paste of the code.

The large amount of figures presented in the text is a sign of the good quality of the work; however, it sometimes introduces limitations: the author does not explain some of the many results presented in the figures, e.g., why does Figure 6.48 (third line of figures) not give values between 40 and 100%?, or

why a 1,5 dBm difference between 2,4 and 5 GHz noise floor is observed in Figure 3.1? The author has not tried to give a reason for this difference. Lastly, the author does not always provide clear conclusions to the results presented and confines to describing the results (e.g., Figure 6.18). In my opinion, the author could reduce the number of figures without losing the main information. Similarly, the format of the figures is not optimal (too small figures), and it is impossible to appreciate some details. For example, in Figure 4.15, some green circles should appear (as stated by the author in the text, page 56), but these circles are not observable in the figures.
Another problem with smooth reading is that there are some grammar errors in the text, and sometimes, the text is a little too verbose and annoying. For example, it is difficult to understand what the author wants to say in the sentence: "The main objective of this chapter is to present recent development and usage possibilities of publicly available real-life diversified data sources as the cornerstone of a novel urban WSN-related research approach".

4. Knowledge of the candidate

The description of the content of the dissertation's chapters is in Section 2 of this Opinion. In the dissertation, the Ph.D. candidate has demonstrated a high knowledge of the matter of his research in the discipline of Information and Communication Technology.
The dissertation contains 214 references, all of them cited in the text of the dissertation and all of them available on the Internet. Regretfully, the references do not include DOI identifiers, and it is unclear when the web pages cited in the references were accessed for the last time. The 214 references are necessary for offering a complete framework of the vast theme, which is the urban wireless network. The number of references is high for a Ph.D. dissertation. This shows that the author has put much effort into analysing the SotA and locating their own ideas well in the framework of similar research efforts.

I consider that the candidate has solid knowledge in his field, mainly due to the long extension of the candidate's involvement in this research. He has been working on ad-hoc wireless networks for more than ten years. However, the results presented in this dissertation relate to his work in the last 4 or 5 years.
Concretely, the results of the work have been presented in 4 publications: two of them in MDPI Sensors and one in MDPI Electronics. One more publication has appeared in the Journal of Telecommunications and Information Technology. The journals are in Q3 or Q4.
Previously, he had published in several other journals with even higher Impact Factor: International Journal of Grid and Distributed Computing, Telecommunications Systems, International Journal of Image Processing & Communications, etc.
The candidate's H-Index (Hirsch) is not available in Google Scholar and Web of Science, at least I could not find them.

5. Conclusion

Taking into account what I have presented above and the requirements imposed by Article 13 of *the Act of 14 March 2003 of the Polish Parliament on the Academic Degrees and the Academic Title* (with amendments)[1], my evaluation of the dissertation according to the three basic criteria is the following:

[1] http://www.nauka.gov.pl/g2/oryginal/2013_05/b26ba540a5785d48bee41aec63403b2c.pdf

A. Does the dissertation present an original solution to a scientific problem? (the selected option is marked with **X**)

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

B. After reading the dissertation, would you agree that the candidate has general theoretical knowledge and understanding of the discipline of **Information and Communication Technology**, and particularly the area of **Delay-tolerant Wireless Networks**?

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

C. Does the dissertation support the claim that the candidate is able to conduct scientific work?

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

Jordi Mongay Batalla

Signature