

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

3/2019

Comparison of Multicast Algorithm Evaluation Results in Low and High Multicast Saturation Environments

K. Stachowiak and P. Zwierzykowski

Paper

3

Spline-Extrapolation Method in Traffic Forecasting in 5G Networks

I. Strelkovskaya, I. Solovskaya, and A. Makoganiuk

Paper

8

Multifactor Authentication and Key Management Protocol for WSN-assisted IoT Communication

T. Ara and M. Prabhakar

Paper

17

A Novel Graph-modification Technique for User Privacy-preserving on Social Networks

S. H. Erfani and R. Mortazavi

Paper

27

An Inclusive Performance Analysis of Single-branch Single-relay AF Transmission in a Mixed Rayleigh-Nakagami- m Fading Environment

S. Benmahmoud

Paper

39

New OSTBC for Blind Channel Estimation and Tracking in MIMO-OFDM Systems

A. Tami, M. Keche, and B. S. Bouazza

Paper

49

Complex Signal Mapping for Improving Spectral Efficiency of Li-Fi Systems

I. K. Sileh, A. A. Abdulkafi, M. K. Hussein, and S. M. Hardan

Paper

58

(Contents Continued on Back Cover)

Editorial Board

Editor-in Chief: ***Adrian Kliks***
Managing Editor: ***Robert Magdziak***
Technical Editor: ***Ewa Kapuściarek***

Editorial Advisory Board

Chairman: ***Andrzej Jajszczyk***
Hovik Baghdasaryan
Naveen Chilamkurti
Luis M. Correia
Luca De Nardis
Nikolaos Dimitriou
Ciprian Dobre
Filip Idzikowski
Albert Levi
Marian Marciniak
George Mastorakis
Constandinos Mavromoustakis
Jordi Mongay Batalla
Klaus Mößner
Imran Muhammad
Mjumo Mzyece
Daniel Negru
Ewa Orłowska
Jordi Perez-Romero
Michał Pióro
Konstantinos Psannis
Salvatore Signorelli
Adam Wolisz
Tadeusz A. Wysocki

ISSN 1509-4553 on-line: ISSN 1899-8852

© Copyright by National Institute of Telecommunications, Warsaw 2019

Circulation: 300 copies

Sowa – Druk na życzenie, www.sowadruk.pl, tel. 22 431-81-40



**Ministry of Science
and Higher Education**
Republic of Poland

Improvement of the linguistic quality of the journal; Assignment of DOI numbers; Access to the antiplagiarism system; Editing and printing of the journal – tasks financed under 706/P-DUN/2019 agreement from budget of the Ministry of Science and Higher Education under the science dissemination fund.

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

I am very delighted to present the next issue – third in 2019 – of the *Journal of Telecommunications and Information Technology* quarterly. This issue is the first to appear under my editorship – Adrian Kliks, affiliated to Poznan University of Technology. It is a great honor for me to serve as Editor in Chief for such a prominent international journal dealing with the recent achievements and trends in the broad ICT domain. At the same time, serving this role is a great and challenging commitment, as the minimum target is to at least maintain the high quality of all previous issues of *JTIT*. I would like to thank the previous Editors Prof. Paweł Szczepański, Krzysztof Borzycki and Marek Jaworski, all members of the Editorial Board, Handling Editors and Reviewers for their enormous work on shaping the high quality of this journal!

As the *JTIT* is an open access journal, we would like to gain more benefits from this great opportunity in the future. The review process will be kept unchanged, i.e., only the best papers that achieve a positive note in the rigorous, double-blind reviewing process can be accepted. Our goal as the Editorial Board is to minimize the review time, so the notification of the decision on each paper is delivered at the earliest possible time. I do hope that it will further enhance the reputation of the journal. If you consider being involved in the reviewing process, please do not hesitate to contact me!

The members of the Editorial Board, as well as the Reviewers, represent various countries from almost all continents, guaranteeing transparency and high quality of the entire review process. Moreover, as the journal is international in its scope and open-access in its nature, it covers up-to-date topics investigated in different places in the world, strengthening the opportunities for prospective collaboration even between distant research centers. I hope that due to this *JTIT* will be the space for exchange of results and ideas between scientists and experts both from academia and industry.

JTIT is indexed in numerous databases, including such prominent ones as Scopus, Index Copernicus, BazTech, EBSCO, INSPEC or Google Scholar, just to mention a few of them. Of course, we are planning to further expand the visibility of our Journal.

Now, I think it is time to invite you to investigate deeply the ten, carefully selected papers collected in this issue. They cover selected research problems in the field of physical, medium access and network layers of communication systems, as well as deal with security issues. We hope you will enjoy this issue.

Adrian Kliks
Editor in Chief

Comparison of Multicast Algorithm Evaluation Results in Low and High Multicast Saturation Environments

Krzysztof Stachowiak and Piotr Zwierzykowski

Faculty of Electronics and Telecommunications, Poznan University of Technology, Poznań, Poland

<https://doi.org/10.26636/jtit.2019.135019>

Abstract—The multicast quality of service-enabled routing is a computationally challenging task. Despite ongoing research efforts, the associated mathematical problems are still considered to be NP-hard. In certain applications, computational complexity of finding the optimal connection between a set of network devices may be a particularly difficult challenge. For example, connecting a small group of participants of a teleconference is not much more complex than setting up a set of mutual point-to-point connections. On the other hand, satisfying the demand for such services as IPTV, with their receivers constituting the majority of the network, requires applying appropriate optimization methods in order to ensure real system execution. In this paper, algorithms solving this class of problems are considered. The notion of multicast saturation is introduced to measure the amount of multicast participants relative to the entire network, and the efficiency of the analyzed algorithms is evaluated for different saturation degrees.

Keywords—*quality-of-service, multicast, routing, multicast saturation.*

1. Introduction

Modern telecommunications networks are tasked with ensuring an optimal transmission of data between remote peers. The methods for optimal execution of this task have been studied by network researchers and engineers for many years now. Routing-related problems have been viewed from the point of view of protocols [1]–[6], as well as algorithms [7]–[12].

To achieve optimal routing, algorithms such as the Dijkstra algorithm are used, for example in the OSPF protocol [13]. In most cases, the optimal path is the one of the lowest cost, though the development and the increasing complexity of the network traffic changes the meaning of optimality in this context. The minimal cost (defined, for example as the number of interconnecting links), has been accompanied by such metrics as point-to-point transmission delay or path reliability [14]. A single network may be used for sending multimedia data (e.g. VOD or YouTube streams), real time data (e.g. packets conveying Skype or Voice-over-IP traffic), alongside classic data for which the requirements have not changed significantly since the introduction of telecommu-

nications networks many years ago. Therefore, engineers often reach out for a multiple criterion evaluation of selected paths or trees (also referred to as multi-criterion or QoS – Quality of Service [15]), e.g. the DUAL algorithm in the EIGRP protocol [16].

Finding optimal routes for multiple communication participants, rather than for a pair of peers, increases the complexity of the problem even further. The number of connected nodes is especially important when compared to the total number of nodes in the network. For example, interconnection of a small group of participants of a teleconference does not require multicast optimization, as the potential gain, compared to setting up individual point-to-point connections, is negligible. On the other hand, for applications involving IPTV, it is not feasible to utilize point-to-point connections between the transmission source and all individual receivers, as they will likely constitute the majority of the network's nodes.

Therefore, it is important to evaluate the performance of routing algorithms with regard to the amount of multicast participants, compared to the number of nodes in the network – the so-called multicast saturation. Based on that statement, a selection of multi-criterion multicast algorithms (solving the MCMST – multi-constrained minimal Steiner tree problem) were compared for different multicast saturation values, and the results are presented in this article.

The article is divided into the following parts. Section 2 presents the mathematical model of the network and the routing problem considered. Section 3 describes MCMST algorithms that were compared in the article. In Section 4, the experiment methodology and results are discussed, while Section 5 concludes the article.

2. Mathematical Model

The network is modeled by an undirected graph $G(E, V)$, where V is a finite set of nodes and $E \subseteq (u, v) : u, v \in V$ is a set of edges that represent point-to-point links. Each of the edges is assigned with M metrics, given by the functions: $(m_i : E \rightarrow \mathbb{R}^+, i = 0, 1, \dots, M-1)$, that reflect additive costs of the relevant edges.

Path $p(s, d)$ from node s to the node d , where $s, d \in V$, is defined as a sequence of non-repeated nodes $v_1, v_2, \dots, v_k \in V$ such that for each $1 \leq i < k$ an edge $(v_i, v_{i+1}) \in E$ and $v_1 \equiv s, v_k \equiv d$. We define the accumulated metrics for the paths, so that the cost of a path p , based on the edges that form it $e \in p \subseteq E$, the i -th is defined as $m_i(p) = \sum_{e \in p} m_i(e)$.

A rooted multicast tree $t(s, d_1, d_2, \dots)$, connecting the source node $s \in V$ with multiple destinations $d_1, d_2, \dots \in D \subseteq V$, is defined as a tree in G , of which the only leaf nodes are the ones from the set $\{s\} \cup D$, with one of them, node s , arbitrarily selected as the root. We define the accumulated cost of tree t correspondingly to the accumulated path's cost, as $m_i(t) = \sum_{e \in t} m_i(e)$. Let

$T(s, d_1, d_2, \dots)$ define the set of all the trees spanning the nodes from set $\{s\} \cup D$. For tree t , we define path $p_t(s, d_i)$ that is connecting nodes s and d_i within the given tree. We define set of constraints C as $(c_i \in \mathbb{R}^+, i = 1, 2, \dots, M-1)$. The constraints are associated with the metrics of the same indices.

The MCMST problem is defined as finding the tree t^* spanning source node s and destination node D that fulfills the following conditions:

$$\begin{aligned} \forall t \in T(s, d_1, d_2, \dots) : m_0(t^*) &\leq m_0(t), \\ \forall d_i \in D, c_j \in C : m_j(p_t(s, d_i)) &\leq c_j. \end{aligned}$$

The multicast saturation is defined as $S = \frac{|\{s\} \cup D|}{|V|}$.

3. Compared Algorithms

In the experiment presented in this article, three different multi-constrained multicast algorithms have been used. Two of them are developed by the author: aggregated MLARAC and RDP, while the third, HMC MC, has been chosen from the literature as a high quality, well-performing algorithm solving the same class of problems.

3.1. Aggregated MLARAC Algorithm

The aggregated MLARAC algorithm [17] is based on the technique presented in [18], [19], solving the multi-criterion multicast routing problem by aggregating solutions to a multi-criterion unicast routing problem between a given source and each individual multicast receiver.

The unicast routing algorithm used in aggregated MLARAC is the MLARAC algorithm, presented in [20]. It is based on the Lagrangian relaxation-based technique from [21], extended to multiple constraints by the author. The original algorithm, reduces the routing problem, by means of the Lagrangian relaxation, to finding an optimal coefficient for a modified cost function that is, in turn, used in the Dijkstra algorithm. Iterative approach is applied to obtain increasingly better approximations of the coefficient, until the optimal value is found.

In the variant proposed by the author, multiple optimization criteria may be applied, which results in the necessity

of finding multiple optimal coefficients. While intersections of single-argument linear functions are used in the original algorithm to obtain the single optimal coefficient, the author presents a method based on finding multi-dimensional hyper-plane intersections to obtain the set of optimal coefficients for the solution of the multi-criterion problem.

3.2. RDP Algorithm

The RDP algorithm is based on simulating multiple Dijkstra algorithm instances (known as convergence processes) concurrently [22]. The processes are performed concurrently and independently, except for the order in which they progress. Information about which nodes have been visited by particular processes is stored by the algorithm core, along predecessor-related information. The selection of the next process is made based on the assumption that labels assigned by the processes to the nodes may be interpreted as the time that has passed in a given process. Once a given node has been visited by all convergence processes, it may be considered the center of a multicast tree – the RDP. There are two important aspects of this approach.

First of all, the Dijkstra algorithm may be enhanced with a custom cost function. Two different cost functions have been used in RDP: first, minimizing the metric m_0 only, and second, aggregating all metrics considered. The aggregation (found in [23]), takes into account metrics m_i , and constraints c_i in the following formula:

$$m_{aggr}(t) = \max \left\{ \frac{m_1(t)}{c_1}, \frac{m_2(t)}{c_2}, \dots \right\}.$$

Secondly, once a RDP candidate has been found, different decisions may be made. In one variant of the algorithm, all subsequent candidates are considered for the final solution. The first to meet the assumed constraints is taken as the result. In a heuristic variant of the algorithm, only the first RDP is considered as the potential solution. If it meets the constraints, it is taken as the ultimate result.

3.3. HMC MC

The Heuristic Multi-Constrained MultiCast algorithm [24] (HMC MC) consists of two phases, with the first of them consisting in finding a partially feasible tree. An attempt is made at building the entire tree, starting from the multicast source, using a modified Dijkstra algorithm that heuristically considers all the criteria – the optimized one, as well as the constraints. It utilizes the same modification as the RDP algorithm. It is possible that the result of the first phase connects all of the multicast participants with a feasible tree, in which case the successful result is returned. However, if any of the multicast receivers has not been feasibly connected with the source, the second phase is performed.

In the second phase, all of multicast participants which have not yet been connected are subject to the correction

procedure. The procedure consists in an attempt to solve an MCOP problem from a given receiver to the source with an associated algorithm – H_MCP. The sub-algorithm is a precomputation-based variant of the unicast point-to-point routing algorithm H_MCOP [25]. The result of this step – a path from the source to one of the receivers – is included in the result, whether it is feasible with regard to the constraints or not.

4. Simulation Experiment

In order to evaluate and compare the aforementioned algorithms, the following experiment was performed. A set of topological structures of 500 nodes was generated using the BRITE generator [26] and the Waxman method [7]. This generator has been used extensively in prior work and has proven to be a valid choice for this type of experiments [17], [20], [22], [27]–[29].

Each of the graph edges was assigned with three metrics: m_0 , m_1 and m_2 . In each experiment iteration, a random group of nodes was selected in a given graph, and the routing problem for fixed constraints was solved using each of the algorithms. The assumed constraints were 17 and 12 for m_1 and m_2 , respectively.

Two rounds of experiments were conducted, for different multicast saturations. First, for the low saturation, assumed sets of 4, 8, 12, 16, 20, 24, 28 multicast participants. In the second round, for the high saturation, sets of 120, 190, 260, 330, 400 multicast participants were considered.

It is important to note that the optimal value of metric m_0 is the lowest, while it is considered that for the constrained metrics m_1 and m_2 , the optimal value is the highest. This stems from the definition of the solved problem. The optimal solution is the one that is of the lowest m_0 metric, while remaining within certain constraints against metrics m_1 and m_2 . Therefore, the result which satisfies the constraints with the highest possible values has the advantage of not acquiring more network resources than necessary.

4.1. Low Multicast Saturation

In Fig. 1, the results of tree metric m_0 are presented as a function of group size, for lower multicast saturation. One may notice that the worst results have been obtained for the RDP algorithm, while aggregated MLARAC and HMC MC algorithms have led to obtaining comparable, better values. Figure 2 demonstrates values of the path metric m_2 under the same simulation conditions. The best results, from the point of view of satisfying the constraints at the lowest cost, were obtained with the use of the RDP algorithm. The second best results were obtained with the aggregated MLARAC algorithm, and the worst m_2 metric results were found using HMC MC.

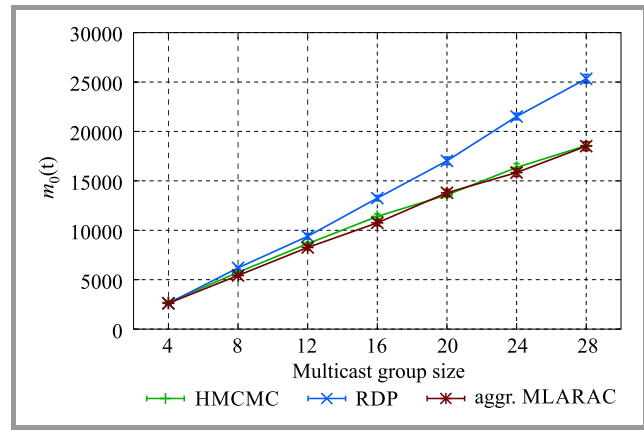


Fig. 1. Average m_0 tree metric values as a function of the multicast group size, for low multicast saturation.

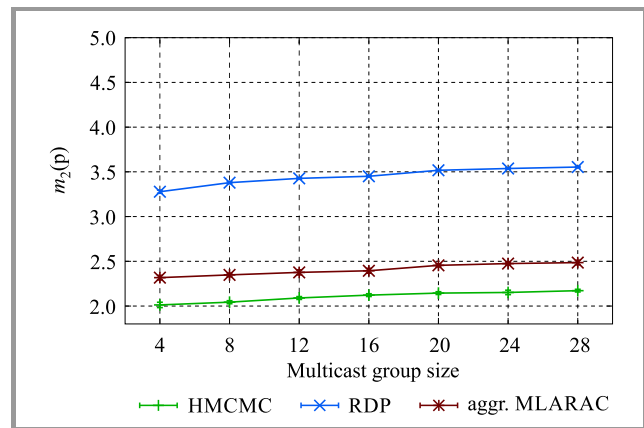


Fig. 2. Average m_2 path metric values as a function of the multicast group size, for low multicast saturation.

4.2. High Multicast Saturation

The following figures show corresponding results for high multicast saturation.

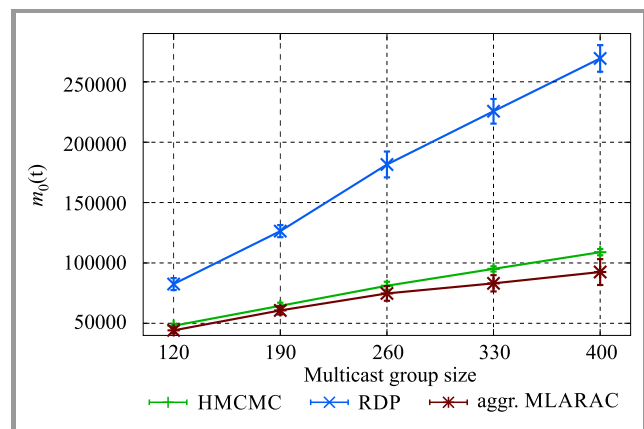


Fig. 3. Average m_0 tree metric values as a function of the multicast group size, for high multicast saturation.

Figure 3 presents the tree metric m_0 for higher multicast saturation values, whereas Fig. 4 demonstrates the tree met-

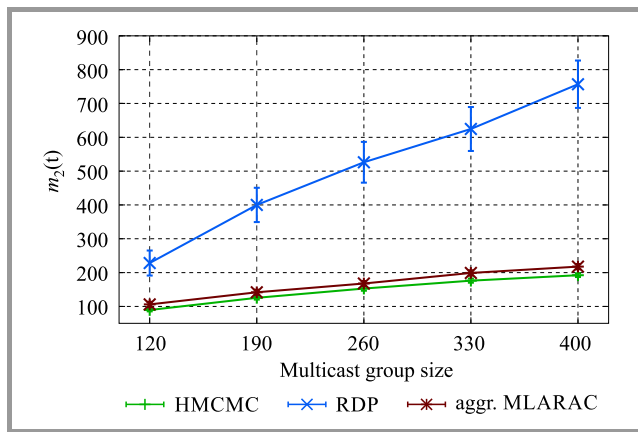


Fig. 4. Average m_2 tree metric values as a function of the multicast group size, for high multicast saturation.

ric m_2 . One may notice how the relationship presented in Figs. 1 and 2 has been maintained under higher multicast saturation conditions.

Such a result is not surprising, since the stochastic character of the generated networks should usually lead to obtaining similar results at larger scales. On the other hand, anomalies may occur leading to different results for different saturation levels. Therefore, it is vital to perform experiments under higher multicast saturation conditions, in order to further support the conclusions that may be drawn from the results obtained for lower values of this particular parameter.

5. Conclusions

The relationship between the number of multicast participants and the network nodes may vary depending on the specific application context. It is important to take this factor into account when suggesting an algorithm for real life use, because different results may be obtained in different operational environments.

The results presented in the article show that the conclusions regarding performance of the presented algorithms, as published in previous works for low multicast saturation levels, still hold in the conditions of higher multicast saturation.

References

- [1] T. Bartczak and P. Zwierzykowski, "Lightweight PIM – a new multicast routing protocol", *Int. J. of Commun. Syst.*, vol. 27, no. 10, pp. 1441–1458, 2014 (doi: 10.12/dac.2407).
- [2] M. Pushpalatha, T. K. Ramesh, C. V. Giriraja, and S. K. Konda, "Power efficient multicast routing protocol for dynamic intra cluster device to device communication", in *Proc. Int. Conf. on Elec., Electron., and Optimiz. Techniq. ICEEOT 2016*, Chennai, India, 2016, pp. 1853–1856 (doi: 10.1109/ICEEOT.2016.7755009).
- [3] V. Matre and R. Karandikar, "Multipath routing protocol for mobile adhoc networks", in *Proc. Sym. on Colossal Data Anal. and Netw. CDAN 2016*, Indore, India, 2016, pp. 428–432 (doi: 10.1109/CDAN.2016.7570946).

- [4] M. Piechowiak, P. Zwierzykowski, P. Owczarek, and M. Wasłowicz, "Comparative analysis of routing protocols for wireless mesh networks", in *Proc. 10th Int. Symp. on Commun. Syst., Netw. and Digit. Sig. Process. CSNDSP 2016*, Prague, Czech Republic, 2016, pp. 1–5 (doi: 10.1109/CSNDSP.2016.7573902).
- [5] T. Bartczak and P. Zwierzykowski, "Performance evaluation of source-specific multicast routing protocols for IP networks", in *Proc. 8th Int. Symp. on Commun. Syst., Netw. & Digit. Sig. Process. CSNDSP 2012*, Poznań, Poland, 2012, pp. 1–6 (doi: 10.1109/CSNDSP.2012.6292693).
- [6] S. Yahiaoui, M. Omar, A. Bouabdallah, E. Natalizio, and Y. Challal, "An energy efficient and QoS aware routing protocol for wireless sensor and actuator networks", *AEU – Int. J. of Electron. and Commun.*, vol. 83, no. Supplement C, pp. 193–203, 2018 (doi: 10.1016/j.aeu.2017.08.045 [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S143484111730897X>).
- [7] B. Waxman, "Routing of multipoint connections", *IEEE J. on Sel. Areas in Commun.*, vol. 6, no. 9, p. 1617–1622, 1988 (doi: 10.1109/49.12889).
- [8] M. Piechowiak, P. Zwierzykowski, K. Stachowiak, and T. Bartczak, "Quality of multicast trees in ad-hoc networks with topology control", in *Proc. 9th Int. Symp. on Commun. Syst., Netw. Digit. Sign CSNDSP 2014*, Manchester, UK, 2014, pp. 7–11 (doi: 10.1109/CSNDSP.2014.6923788).
- [9] Z. Xiaoxia, S. Xuanyao, T. Jiwei, and L. Guoxuan, "A hybrid ant colony optimization based on cloud model for OoS multicast routing problem", in *Proc. 29th Chinese Control and Dec. Conf. CCDC 2017*, Chongqing, China, 2017, pp. 6551–6554 (doi: 10.1109/CCDC.2017.7978352).
- [10] N. Anwar and H. Deng, "Ant colony optimization based multicast routing algorithm for mobile ad hoc networks", in *Proc. Adv. in Wirel. and Optical Commun. RTUWO 2015*, Riga, Latvia, 2015, pp. 62–67 (doi: 10.1109/RTUWO.2015.7365721).
- [11] M. Mahseur, Y. Meraihi, A. Boukra, and A. Ramdane-Cherif, "QoS multicast routing based on a hybrid quantum evolutionary algorithm with firefly algorithm", in *Proc. 5th Int. Conf. on Elec. Engin. – Boumerdes ICEE-B 2017*, Boumerdes, Algeria, 2017, pp. 1–6 (doi: 10.1109/CEE-B.2017.8192154).
- [12] D. Papadimitriou, B. Fortz, and E. Gorgone, "Lagrangian relaxation for the time-dependent combined network design and routing problem", in *Proc. IEEE Int. Conf. on Commun. ICC 2015*, London, UK, 2015, pp. 6030–6036 (doi: 10.1109/ICC.2015.7249283).
- [13] J. Moy, "OSPF Version 2", RFC 2328 (Standard), IETF, Apr. 1998, updated by RFC 5709 [Online]. Available: <http://www.ietf.org/rfc/rfc2328.txt>
- [14] J. F. Kurose and K. Ross, *Computer Networking: A Top-Down Approach Featuring the Internet*, 2nd ed. Boston, MA, USA: Addison-Wesley, 2003 (ISBN: 0201976994).
- [15] S. Chen and K. Nahrstedt, "An overview of quality of service routing for next-generation high-speed networks: problems and solutions", *IEEE Netw.*, vol. 12, no. 6, pp. 64–79, 1998 (doi: 10.1109/65.752646).
- [16] D. Savage, J. Ng, S. Moore, D. Slice, P. Paluch, and R. White, "Cisco's Enhanced Interior Gateway Routing Protocol (EIGRP)", RFC 7868 (Informational), RFC Editor, Fremont, CA, USA, pp. 1–80, May 2016 [Online]. Available: <https://www.rfc-editor.org/rfc/rfc7868.txt>
- [17] K. Stachowiak, P. Zwierzykowski, M. Piechowiak, and P. Prokopowicz, "Evaluation of different path composition approaches to multicast optimization with fuzzy set", *Mediterranean J. of Electron. and Commun.*, vol. 9, no. 3, pp. 594–600, 2013.
- [18] M. Piechowiak, P. Zwierzykowski, and M. Stasiak, "Multicast routing algorithm for packet networks with the application of the Lagrange relaxation", in *Proc. of the 14th Int. Telecommun. Netw. Strat. and Plann. Symp. NETWORKS 2010*, Warsaw, Poland, 2010, pp. 197–202 (doi: 10.1109/NETWKS.2010.5624921).
- [19] Y. Xiao, K. Thulasiraman, G. Xue, and M. Yadav, "QoS routing under multiple additive constraints: A generalization of the Larac algorithm", *IEEE Trans. on Emerg. Topics in Comput.*, vol. 4, no. 2, pp. 242–251, 2016 (doi: 10.1109/TETC.2015.2428654).

- [20] K. Stachowiak, J. Weissenberg, and P. Zwierzykowski, "Lagrangian relaxation in the multicriterial routing", in *Proc. IEEE AFRICON 2011*, Livingstone, Zambia, 2011, p. 646–651 (doi: 10.1109/AFRCON.2011.6072070).
- [21] A. Juttner, B. Szviatovski, I. Mecs, and Z. Rajko, "Lagrange relaxation based method for the QoS routing problem", in *Proc. IEEE INFOCOM 2001 – Twentieth Ann. Joint Conf. of the IEEE Comp. and Commun. Soc.*, Anchorage, AK, USA, 2001, vol. 2, pp. 859–868 (doi: 10.1109/INFCOM.2001.916277).
- [22] K. Stachowiak and P. Zwierzykowski, "Rendezvous point based approach to the multi-constrained multicast routing problem", *AEU – Int. J. of Electron. and Commun.*, vol. 68, no. 6, pp. 561–564, 2014 (doi: 10.1016/j.aeue.2014.01.002).
- [23] H. D. Neve and P. V. Mieghem, "Tamcra: a tunable accuracy multiple constraints routing algorithm", *Comp. Commun.*, vol. 23, no. 7, pp. 667–679, 2000 (doi: 10.1016/S0140-3664(99)00225-X).
- [24] F. Gang, "A multi-constrained multicast QoS routing algorithm", *Comp. Commun.*, vol. 29, no. 10, p. 1811–1822, 2006 (doi: 10.1016/j.comcom.2005.10.014).
- [25] T. Korkmaz and M. Krunz, "Multi-constrained optimal path selection", in *Proc. of IEEE Int. Conf. on Comp. Commun. INFOCOM 2001*, Anchorage, AK, USA, 2001, pp. 834–843 (doi: 10.1109/INFCOM.2001.916274).
- [26] A. Medina, A. Lakhina, I. Matta, and J. Byers, "BRIT: An approach to universal topology generation", in *Proc. 9th Int. Symp. on Model., Anal. and Simul. of Comp. and Telecommun. Syst. MASCTOS 2001*, Cincinnati, OH, USA, 2001, pp. 346–356 (doi: 10.1109/MASCTOS.2001.948886).
- [27] K. Stachowiak and P. Zwierzykowski, "Lagrangian relaxation and linear intersection based QoS routing algorithm", *Int. J. of Electron. and Telecommun.*, vol. 58, no. 4, p. 307–314, 2013 (doi: 10.2478/v10177-012-0042-2).
- [28] M. Piechowiak, K. Stachowiak, and P. Zwierzykowski, "Properties of multiconstrained multicast trees in ad-hoc networks with topology control", in *Proc. of the IEICE Inform. and Commun. Technol. Forum ICTF 2015*, Manchester, UK, 2015, pp. 1–5 2015.
- [29] M. Piechowiak and S. Pucek, "Topology properties of ad-hoc networks", in *Image Processing and Communications Challenges 4*, R. S. Choraś, Ed. Berlin, Heidelberg: Springer, 2013, pp. 233–240 (doi: 10.1007/978-3-642-32384-3_29).



Krzysztof Stachowiak received his Ph.D. degree in Telecommunications from Poznan University of Technology, Poland, in 2018. He is involved in the research regarding multicast routing in packet switching networks, mainly consisting in the analysis of existing QoS routing algorithms, as well as in

inventing new proposals. The main subject of his research is multicriterial optimization which has resulted in a series of papers on the subject of the linear and non-linear Lagrangian relaxation. Krzysztof Stachowiak has participated in organizing two international scientific conferences, and has been a participant of several others events of this type. Besides conducting scientific research, he is also a professional programmer in the domain of embedded software development.

 <https://orcid.org/0000-0003-4993-9820>

E-mail: krzysiek.stachowiak@gmail.com

Chair of Communication and Computer Networks

Faculty of Electronics and Telecommunications

Poznan University of Technology

Pl. Marii Skłodowskiej-Curie 5

60-965 Poznań, Poland



Piotr Zwierzykowski received his M.Sc. and Ph.D. degrees in Telecommunications from Poznan University of Technology, Poland, in 1995 and 2002, respectively. Since 1995 he has been working at the Faculty of Electronics and Telecommunications, Poznan University of Technology. He is currently an Assistant Professor at the Chair

of Communications and Computer Networks. He is the author, or co-author, of over 200 papers and three books. Dr Piotr Zwierzykowski is engaged in conducting research and teaching activities in the area of computer networks, multicast routing algorithms and protocols, as well as performance analysis of multiservice switching systems. Recently, his has been focused primarily on modeling multiservice cellular networks.

 <https://orcid.org/0000-0001-5609-1026>

E-mail: piotr.zwierzykowski@put.poznan.pl

Chair of Communication and Computer Networks

Faculty of Electronics and Telecommunications

Poznan University of Technology

Pl. Marii Skłodowskiej-Curie 5

60-965 Poznań, Poland

Spline-Extrapolation Method in Traffic Forecasting in 5G Networks

Irina Strelkovskaya, Irina Solovskaya, and Anastasiya Makoganiuk

O. S. Popov Odessa National Academy of Telecommunications, Odessa, Ukraine

<https://doi.org/10.26636/jtit.2019.134719>

Abstract—This paper considers the problem of predicting self-similar traffic with a significant number of pulsations and the property of long-term dependence, using various spline functions. The research work focused on the process of modeling self-similar traffic handled in a mobile network. A spline-extrapolation method based on various spline functions (linear, cubic and cubic B-splines) is proposed to predict self-similar traffic outside the period of time in which packet data transmission occurs. Extrapolation of traffic for short- and long-term forecasts is considered. Comparison of the results of the prediction of self-similar traffic using various spline functions has shown that the accuracy of the forecast can be improved through the use of cubic B-splines. The results allow to conclude that it is advisable to use spline extrapolation in predicting self-similar traffic, thereby recommending this method for use in practice in solving traffic prediction-related problems.

Keywords—*quality of service, self-similar traffic, spline functions, error of recovery.*

1. Introduction

The development of fifth-generation 5G/NR (New Radio) mobile communication networks is associated with the rapid growth in the number of high-speed services offered to users. These include enhanced Mobile Broadband (eMBB) access, massive Machine Type Communications (mMTC) and Ultra-reliable and Low Latency Communications (URLLC), as defined in the recommendations of the 3GPP technical report 38.913 [1], [2]. These services require increased data transmission speeds and quality of service levels, perceived as an optimal combination of packet delay time, packet loss probability and throughput. The main difficulty in solving this problem consists in significant and frequent bursts of intensity, which appear to be statistically similar at different time scales. This property allows us to predict the occurrence of this condition in the future, based on the accumulated statistical data or based on the results of traffic modeling [3]–[6]. In practice, the appearance of a significant number of long-term traffic intensity ripples at arbitrary points in time often leads to a sharp increase in packet delay time, which causes overloading of

network nodes and buffer devices and, accordingly, has a significant impact on traffic.

Predicting self-similar traffic will enable to handle the potential peak loads in the network and to manage traffic in an efficient manner, thereby ensuring the required QoS. Given the above facts, the urgency of the problem of predicting self-similar traffic is obvious.

The purpose of the work is to solve the problem of predicting self-similar traffic and choosing the method by which traffic will be restored more precisely beyond the limits of its determining time intervals.

2. Problem Statement

Function extrapolation is a continuation of a function outside its domain of definition, in which the continued function belongs to a given class. Extrapolation is usually performed using formulas that rely on information about the function's behavior at a certain set of points (extrapolation nodes) belonging to its domain of definition [7]. Sometimes, when extrapolating functions, not the entire domain of definition is used. Instead, only a part of it is taken into consideration. In fact, extrapolation of the values of judgment of a given function over a specific portion thereof is performed. In this case, extrapolation formulas give, in particular, the values of the function at the corresponding points of its domain of definition. This method is often used when solving practical problems, when there is no sufficient information available that is necessary to determine the values outside the considered part of the function's domain of definition [7].

A properly selected forecasting method allows to obtain predictions about the availability of the necessary bandwidth and delay time for managing network peak loads. The prediction results will help the operator perform traffic management functions, thereby ensuring the required QoS levels [8].

The prediction problems solved by the authors of the past were based on the extrapolation of random processes, performed using the Lagrange interpolation multi-members, Chebyshev polynomials, etc. Some prediction-related issues are considered in [9]–[11], where numerical methods

of Hermite interpolation, as well as trigonometric interpolation and Richardson extrapolation were used. These methods allow solving the forecasting problem with a given level of accuracy. However, the proposed forecasting methods are not universal.

There is no single method for predicting the characteristics of self-similar traffic that would enable real-time prediction of self-similar traffic behaviors, giving the mobile operator the opportunity to take timely measures to prevent network node overloads, and to avoid the corresponding effects of these changes on QoS.

An extrapolation method based on spline functions, which has a number of advantages compared to its already known counterparts, offers the following features [5]:

- splines are more resistant to local perturbations, that is, the behavior of the spline in the vicinity of a point does not affect the behavior of the spline as a whole, as is the case in polynomial interpolation,
- good convergence of spline interpolation in contrast to the polynomial. In particular, for functions with irregular smoothness properties (for example, self-similar traffic) spline interpolation is indisputably high.

Let us consider spline-extrapolation of self-similar traffic, modeled with the use of the Matlab Simulink package for a given set of initial data, using various spline functions (linear, cubic, and B-splines). Spline functions of odd degrees (linear, cubic, cubic B-splines) are used in this paper due to their properties of the minimum norm and the best approximation. This allows to obtain more accurate results in terms of convergence [12].

3. Related Work

A significant number of scientific papers [13]–[19] is devoted to traffic forecasting issues. Their authors propose various forecasting methods. The classical approach to forecasting traffic characteristics is used in [13], where a comparison of various methods, such as the method based on a polynomial extrapolation of Lagrange, Markov chains and the method of automatic and trainees show the feasibility of using an extrapolation method in predicting self-similar traffic. One of the generally accepted traffic prediction mechanisms has the form of linear regression models, such as Autoregressive Integrated Moving Average (ARIMA) and Fractal Autoregressive Integrated Moving Average (FARIMA), as considered in [4]–[16]. These are used only for short-term forecasts, when solving traffic control and network performance problems.

However, their use is based on an accurate estimate of the Hurst parameter which, when predicting traffic in real-time, is fairly difficult to estimate. Therefore, the use of ARIMA and FARIMA models is possible only for traffic with a slight degree of self-similarity.

The neural network prediction method, as discussed in a number of papers, for example in [17], [18], allows to solve a number of practical forecasting problems, such as dynamic bandwidth redistribution, in order to make the optimal use of available network resources and to maintain QoS.

However, it is important to note that the use of the neural network method implies the need for network training. This is a complex and time-consuming process, and even a trained network is not always clearly predictable due to heuristic approaches to training design.

The forecast accuracy in such a network depends on the number of options for its training. The implementation requires high computing capabilities. In [19], the results of a comparison of network traffic predictions made using linear regression models and SARIMA neural networks are shown, indicating that in most cases the use of complicated and labor-intensive techniques of neural networks is impractical.

The methods of extrapolation [9]–[11], [13]–[19] are labor-intensive and suffer from considerable errors under conditions of frequent traffic intensity bursts. It should be noted that they are complex when forecasting in real time. Solving the problem of extrapolation by cubic splines or their piece-wise polynomial representation in many cases is a convenient tool, both for solving theoretical problems and for calculating the terms. However, in a number of applications, it is more efficient to represent cubic splines via B-splines.

Therefore, the development of effective methods of prediction of self-similar traffic, reducing the computational complexity and, thus, providing solutions for networks with significant bandwidths, is an important task.

4. Method of Self-Similar Traffic Prediction

Consider self-similar traffic affecting segment $[a; b]$. Let partition $\Delta: a = x_0, x_1, \dots, x_N = b$ be given in interval $[a; b]$. The first-degree spline $S_1(x)$ on grid Δ is a continuous piece-wise linear function. Let Δ grid points be given the values of self-similar traffic $f_i = f(x_i)$, which describes function $f(x)$, defined for interval $[a; b]$. The interpolation spline is defined by [5], [6], [20]:

$$S_1(x_i) = f_i, \quad i = 0, \dots, N. \quad (1)$$

Geometrically, it is a broken line passing through the points (x_i, y_i) , where $y_i = f(x_i)$, denoted by $h_i = x_{i+1} - x_i$. Then, according to [20], for $x \in [x_i, x_{i+1}]$, $i = 0, \dots, N-1$, the linear spline will have the following form:

$$S_1(x_i) = f_i \frac{x_{i+1} - x}{h_i} + f_{i+1} \frac{x - x_i}{h_i}, \quad (2)$$

or

$$S_1(x_i) = f_i \frac{x - x_i}{h_i} (f_{i+1} - f_i). \quad (3)$$

The authors will also consider a cubic interpolation spline $S_3(x)$, constructed similarly to the linear spline, with the only difference being that this is a cubic function on each interval $[x_i, x_{i+1}]$, $i = 1, \dots, N-1$. According to [5], [6], [20], for $x \in [x_i, x_{i+1}]$, $i = 0, 1, \dots, N-1$ the cubic spline is:

$$S_3(x) = f_i(1-t)^2(1+2t) + f_{i+1}t^2(3-2t) + m_i h_i t(1-t)^2 - m_{i+1} h_i t^2(1-t), \quad (4)$$

where $t = \frac{x-x_i}{h_i}$, $S_3(x_i) = f_i$, $S_3(x_{i+1}) = f_{i+1}$, $m_i = S'(f; x_i)$, or

$$S_3(x) = f_i(1-t) + (1+2t)f_{i+1}t - \frac{h_i^2}{6}(1-t)[(2-t)M_i + (1+t)M_{i+1}], \quad (5)$$

where $S''(x_i) = M_i$, $S''(x_{i+1}) = M_{i+1}$.

The boundary conditions are used to determine the cubic spline of the Eq. (4) in the interval $[a; b]$ [5], [6], [20]:

$$S'(f; a) = f'(a), \quad S'(f; b) = f'(b). \quad (6)$$

For determining a cubic spline type given by Eq. (5), the boundary conditions of the form [5] is used:

$$S''(f; a) = f''(a), \quad S''(f; b) = f''(b). \quad (7)$$

Consider a uniform partition of interval $[a; b]$, then:

$$h_i = h = \frac{b-a}{N}, \quad i = 0, 1, \dots, N-1.$$

Let us construct a cubic B-spline, different from zero in the interval (x_{i-2}, x_{i+2}) . B-splines of odd degrees are conveniently numbered by the middle node of their carrier intervals. The desired B-spline will be denoted by $B_i(x)$. Setting $y_p = B_i(x_p)$, $M_p = B_i''(x_p)$ for $B_i(x)$ we have [20], [21]:

$$\mu_p M_{p-1} + 2M_p + \lambda_p M_{p+1} = \frac{6}{h_{p-1} + h_p} \left(\frac{y_{p+1} - y_p}{h_p} - \frac{y_p - y_{p-1}}{h_{p-1}} \right), \quad (8)$$

where $p = i-1, i, i+1$, $\mu_i = \frac{h_{i-1}}{h_{i-1} + h_i}$, $\lambda_i = 1 - \mu_i$.

As $B_i(x) = 0$ for $x \notin [x_{i-2}, x_{i+2}]$, then:

$$B_i^{(r)}(x_{i-1}) = B_i^{(r)}(x_{i+2}) = 0, \quad r = 0, 1, 2. \quad (9)$$

Taking into account the fact that the B-spline has a ratio [20], [21]:

$$B_i(x) = y_i(1-t) + y_{i+1}t - \frac{h_i^2}{6}t(1-t) \times [(2-t)M_i + (1+t)M_{i+1}], \quad (10)$$

where $x \in [x_i, x_{i+1}]$, $t = \frac{x-x_i}{h_i}$, $h_i = x_{i+1} - x_i$,

we get:

$$B_i''(x) = \frac{y_{i+1} - y_i}{h_i} - \frac{h_i}{6} [(2-6t+3t^2)M_i + (1-3t^2)M_{i+1}], \quad (11)$$

$$B_i''(x) = M_i(1-t) + M_{i+1}t. \quad (12)$$

Then, condition (9), due to Eqs. (10)–(12), may be represented by:

$$\begin{cases} y_{i-2} = y_{i+2} = 0, & M_{i-2} = M_{i+2} = 0, \\ y_{i-1} = \frac{1}{6}h_{i-2}^2 M_{i-1}, & y_{i+1} = \frac{1}{6}h_{i+1}^2 M_{i+1} \end{cases}. \quad (13)$$

The parameters found in Eq. (13) are excluded from Eq. (8), which will lead to [20], [21]:

$$\begin{cases} (h_{i-2} + h_{i-1})(h_{i-2} + 2h_{i-1})M_{i-1} + h_{i-1}^2 M_i = 6y_i, \\ (h_{i-2} + h_{i-1})M_{i-1} + (h_{i-1} + h_i)M_i + (h_i + h_{i+1})M_{i+1} = 0, \\ h_i^2 M_i (h_i + h_{i+1})(2h_i + h_{i+1})M_{i+1} = 6h_i. \end{cases} \quad (14)$$

The result is a system of three equations for finding four parameters: y_i , M_{i-1} , M_i , M_{i+1} .

Assuming:

$$y_i = \frac{h_{i-1}(h_{i-2} + h_{i-1})(2h_i + h_{i+1}) + h_i(h_i + h_{i+1})(h_{i-2} + 2h_{i-1})}{(h_{i-1} + h_i)(h_{i-2} + h_{i-1} + h_i)(h_{i-1} + h_i + h_{i+1})}, \quad (15)$$

from the Eq. (14) we get:

$$\begin{cases} M_{i-1} = \frac{6}{(h_{i-2} + h_{i-1})(h_{i-2} + h_{i-1} + h_i)}, \\ M_i = \frac{6[(h_{i-2} + h_{i-1} + h_i)^{-1} + (h_{i-1} + h_i + h_{i+1})^{-1}]}{h_{i-1} + h_i}, \\ M_{i+1} = \frac{6}{(h_i + h_{i+1})(h_{i-1} + h_i + h_{i+1})}. \end{cases} \quad (16)$$

Equations (13), (15), (16) determine the spline $B_i(x)$ in the interval $[x_{i-2}, x_{i+2}]$.

It is necessary to restore self-similar traffic outside the interval $[a; b]$, namely, to the right of the point b . For definiteness, let this be a point x_c , $x_c > x_N = b$, $x_c - b = h$, where h is a step partition of the interval $[a; b]$.

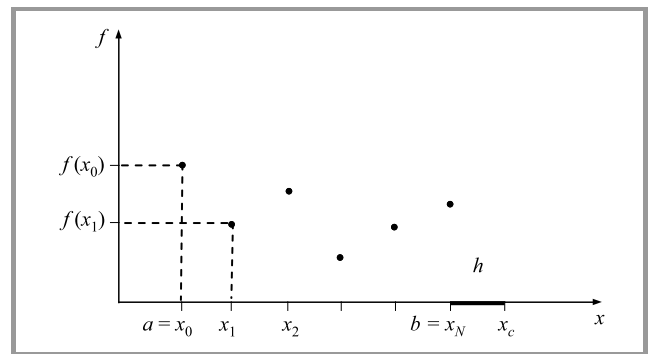


Fig. 1. Extrapolation of self-similar traffic on the interval $[a; b]$ on condition $f(x_c) = f(x_1)$.

Let us now consider short-term and long-term predictions of the characteristics of self-similar traffic using spline extrapolation.

First, a spline function (linear, cubic, or cubic B-spline) on the segment $[b; x_c]$ is constructed with the assumption that $f(x_c) = f(x_1)$. Next, we construct a linear or cubic spline or a cubic B-spline, respectively, on the segment $[b; x_c]$ (Fig. 1). Then, on the segment $[b; x_c]$, a short-term forecast of self-similar traffic is obtained.

In the second case (Fig. 2), the set $f(x_k) = f(x_c)$, with $x_k = x_0 + kh$, where k is natural, is considered. If $kh \neq x_c - b$, then for $f(x_c)$ we take the values of function $f(x)$ which are closest to the point x_k .

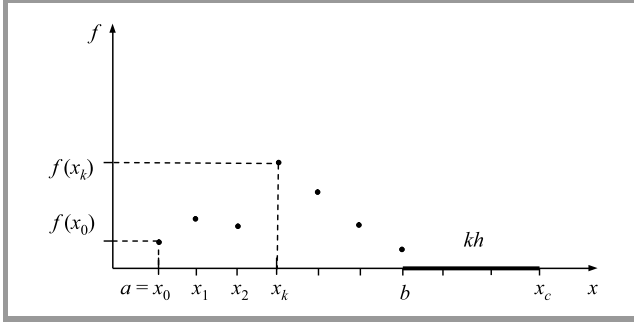


Fig. 2. Extrapolation of self-similar traffic on the interval $[a; b]$ on condition $f(x_{kh}) = f(x_c)$.

According to [20], the error of restoring self-similar traffic on the interval $[b; x_c]$, using the following theorems may be found.

If a spline of the first degree $S_1(x)$ interpolates a continuous function $f(x)$ on the grid Δ , then the estimation of the error is valid:

$$|S_1(x) - f(x)| \leq \omega(f), \quad (17)$$

where $\omega(f)$ – module of a continuous function of the form.

$$\omega(f) = \max_{0 \leq i \leq N-1} \omega(f)' = \max_{0 \leq i \leq N-1} \max_{x', x'' \in [x_i, x_{i+1}]} |f(x'') - f(x')|. \quad (18)$$

If the spline cubic $S_3(x)$ interpolates the continuous function $f(x)$ on the net Δ and satisfies the boundary conditions (6) or (7), then:

$$\|S(x) - f(x)\|_C \leq \left(1 + \frac{3}{4}\rho\right)\omega(f), \quad (19)$$

where:

$$\rho = \frac{\max_i h_i}{\min_i h_i},$$

$$\|f(x)\|_C = \max_{x \in [a, b]} |f(x)|,$$

$C = C[a; b]$ – the function is continuous on the interval $[a; b]$.

Let us find the error of extrapolation of self-similar traffic using cubic B-splines. The interpolation cubic spline $S(x)$ can be found using its B-spline representation [20]:

$$S(x) = \sum_{i=-1}^{N+1} b_i B_i(x). \quad (20)$$

The quality of the interpolation function is characterized as $R(x) = S(x) - f(x)$ and it depends on differential properties of the interpolated function $f(x)$.

Consider a spline satisfying the condition:

$$S(f; x_i) = f_i, \quad i = 0, 1, \dots, N, \quad (21)$$

with a boundary:

$$S'(f; a) = f'(a), \quad S'(f; b) = f'(b). \quad (22)$$

To determine the coefficients b_i , the system of equations is obtained by [20], [21]:

$$\begin{cases} b_{-1}B'_{-1}(x_0) + b_0B'_0(x_0) + b_1B'_1(x_0) = f'_0, \\ b_{i-1}B_{i-1}(x_1) + b_iB_i(x_1) + b_{i+1}B_{i+1}(x_1) = f_i, \\ b_{N-1}B'_{N-1}(x_N) + b_NB'_N(x_N) + b_{N+1}B'_{N+1}(x_N) = f', \end{cases} \quad (23)$$

where $i = 0, \dots, N$.

In the periodic case the Eqs. (17) describing the problem of extrapolation, are:

$$b_{i-1}B_{i-1}(x_1) + b_iB_i(x_i) + b_{i+1}B_{i+1}(x_i) = f_i, \quad i = 1, \dots, N. \quad (24)$$

In matrix form, this can be written as:

$$Ab = f, \quad (25)$$

where $b = (b_1, \dots, b_N)^T$, $f = (f_1, \dots, f_N)^T$ denotes transposition.

The error of extrapolation of self-similar traffic is calculated using Eq. (18)

$$A(b - f) = f - Af. \quad (26)$$

Consider the space $C[a; b]$ to be continuous on $[a; b]$, such as:

$$\|f(x)\|_{C[a, b]} = \max_{x \in [a, b]} |f(x)|.$$

On the grid $\Delta: a = x_0 < x_1 < \dots < x_N = b$, these functions are characterized by their oscillation on segment $[x_i, x_{i+1}]$:

$$\omega_i(f) = \max_{x', x'' \in [x_i, x_{i+1}]} |f(x'') - f(x')|,$$

and:

$$\omega(f) = \max_{0 \leq i \leq N-1} \omega_i(f).$$

If the characteristic of the function is independent of the grid Δ , there is the modulus of $\omega(f; h)$, which is defined as [20]:

$$\omega_i(f; h) = \max_{\substack{x', x'' \in [a, b] \\ |x'' - x'| \leq h}} |f(x'') - f(x')|, \quad h \leq b - a.$$

If $\bar{h} = \max_i h_i$, then the following inequalities are:

$$\omega_i(f) \leq \omega(f) \leq \omega(f; \bar{h}).$$

Then

$$\begin{aligned} \|f - Af\| &= \max_i |f_i - f_{i-1}B_{i-1}(x_i) - f_iB_i(x_i) - f_{i+1}B_{i+1}(x_i)| \\ &\leq \max_i \left\{ B_{i-1}(x_i) |f_i - f_{i-1}| + B_{i+1}(x_i) |f_i - f_{i+1}| \right\} \leq \omega(f). \end{aligned}$$

Using the property of normalized B-splines $\sum_i B_i(x) = 1$ and the Eq. (19) we have:

$$\|b - f\| \leq \|A^{-1}\| \omega(f), \quad (27)$$

and

$$\begin{aligned} \|S(x) - f(x)\| &= \left\| \sum_{i=-1}^{N+1} [b_i - f(x)] B_i(x) \right\| \\ &\leq \left\| \sum_{i=-1}^{N+1} [b_i - f_i] B_i(x) \right\| + \left\| \sum_{i=-1}^{N+1} [f_i - f(x)] B_i(x) \right\|. \end{aligned}$$

Using Eq. (20), the results are obtained by:

$$\left\| \sum_{i=-1}^{N+1} [b_i - f_i] B_i(x) \right\| \leq \|b - f\| \leq \|A^{-1}\| \omega(f).$$

In addition, for any $x \in [x_i, x_{i+1}]$, $i = 0, 1, \dots, N$

$$\left\| \sum_{i=-1}^{N+1} [f_i - f(x)] B_i(x) \right\| \leq \sum_{p=i-1}^{i+2} |f_p - f(x)| b_p(x) \leq 2\omega(f).$$

Finally

$$\|S(x) - f(x)\| \leq (2 + \|A^{-1}\|) \omega(f). \quad (28)$$

According to [20], the matrix A will be ill-conditioned if $\rho \geq \frac{3+\sqrt{5}}{2}$, then it is advisable to use a grid with a partition step for which $\rho < \frac{3+\sqrt{5}}{2}$.

5. Simulations

The spline-extrapolation method described in this paper is used for high-speed video and voice traffic, which are described by the Weibull distribution [3]. The problem concerned may be modeled by other distributions, for example - the Pareto distribution. However, it describes data traffic only.

The authors simulate self-similar traffic for the queuing system (QS) $W_B/M/1/K$ on a 3000–4000 ms segment, where:

- λ is intensity of packet arrival for servicing in the QS, $\lambda = 150$ packet per second,
- μ is packet service duration, $\mu = 125$ packet/s,
 - κ is length of the packet queue, $\kappa = 200$ packets,
 - Hurst parameter $H = 0.75$,
 - Weibull distribution parameters are $\alpha \approx 0.5$ and $\beta \approx 17.321$.

Self-similar traffic simulation results obtained in the Simulink package of the Matlab environment for specific initial data are shown in Fig. 3, where n is the number of packets, t is time of arrival of packets [5], [6].

According to Fig. 3, it can be seen that for the self-similar traffic obtained on the 3000–4000 ms segment, there is large-scale invariance, a significant amount of “bursts” of traffic intensity and a long-term relationship between the moments of their arrival.

Next, consider the extrapolation of traffic on the interval of [3800; 3850] ms and compare it with the results of modeling self-similar traffic.

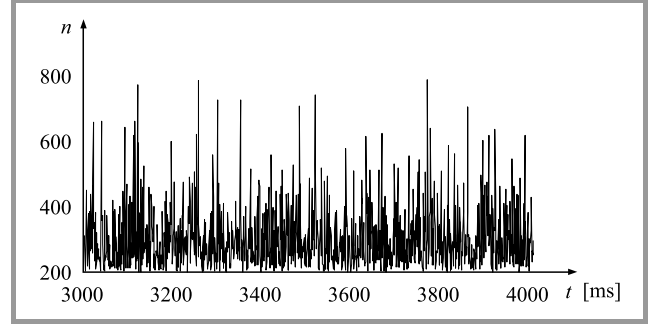


Fig. 3. The results of modeling self-similar traffic.

Using a linear spline for simulated self-similar traffic on the 3800–3850 ms segment, extrapolation of traffic is obtained, which is shown in Fig. 4.

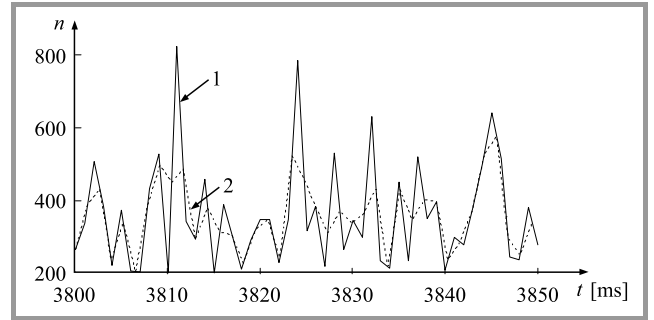


Fig. 4. The results of extrapolating self-similar traffic on the 3800–3850 ms segment using the linear spline function: 1 – self-similar traffic, 2 – extrapolation of self-similar traffic using the linear spline function.

The error in recovering self-similar traffic based on a linear spline will be evaluated according to Eqs. (17)–(18). The calculation results are shown in Table 1.

It is easy to see that the use of linear spline functions is linked with the presence of a significant error, which most often appears on segments where graphs of traffic intensity have periodic “bursts”. It should be noted that the studies were conducted using a small segment of 3800–3850 ms, with step $h = 1$. The split step increase will inevitably lead to an increase in the values of extrapolation error. From this, it follows that the use of spline extrapolations based on linear splines to predict self-similar traffic with significant and frequent bursts of intensity is impractical.

However, during the experiment, self-similar traffic with the Hurst coefficient $H=0.75$ was used, possibly for self-similar traffic with a lower self-similarity value using linear splines. This assumption will be considered in future work.

Table 1
The error of recovery of self-similar traffic
based on a linear spline

Interval	Interval time [ms]	Error value
$[x_0; x_1]$	3800–3801	0.01
$[x_1; x_2]$	3801–3802	2.5
$[x_2; x_3]$	3802–3803	73.3
$[x_3; x_4]$	3803–3804	4.1
...	...	...
$[x_{10}; x_{11}]$	3810–3811	248.2
$[x_{11}; x_{12}]$	3811–3812	335.53
$[x_{12}; x_{13}]$	3812–3813	11.1
$[x_{13}; x_{14}]$	3813–3814	0.1
...	...	...

Using the cubic spline for simulated self-similar traffic on the 3800–3850 ms segment, extrapolation of traffic is obtained, as shown in Fig. 5. The error in recovering self-similar traffic using a cubic spline will be estimated using Eq. (19). The calculation results are summarized in Table 2.

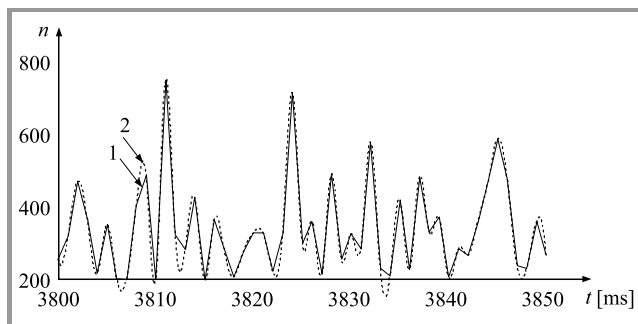


Fig. 5. Extrapolated self-similar traffic on the 3800–3850 ms segment using cubic spline functions: 1 – self-similar traffic, 2 – extrapolation of self-similar traffic using a cubic spline function.

According to the results of the extrapolation of self-similar traffic on the 3800–3850 ms segment using the cubic spline function shown in Fig. 5, the recovery errors are obtained on the segments of the graph, where the traffic intensity has “spikes” at peak points. In general, the use of cubic splines allows you to perform a short-term forecast of traffic parameters and get the predicted “route” of self-similar traffic.

It is possible to increase the accuracy of prediction of traffic characteristics using the spline-extrapolation method based on cubic B-splines. It is known that B-splines have a local character and in their construction several values are used on the considered segments. Therefore, they make it possible to obtain better results compared to cubic splines, the coefficients of which are calculated as a function over the entire domain of the initial function, in this case, at the given interval of 3800–3850 ms.

Table 2
Accuracy of recovery of self-similar traffic
based on cubic spline

Interval	Interval time [ms]	Error value
$[x_0; x_1]$	3800–3801	4.6
$[x_1; x_2]$	3801–3802	0.15
$[x_2; x_3]$	3802–3803	2.44
$[x_3; x_4]$	3803–3804	1.46
...	...	...
$[x_{10}; x_{11}]$	3810–3811	0.12
$[x_{11}; x_{12}]$	3811–3812	2.1
$[x_{12}; x_{13}]$	3812–3813	10.6
$[x_{13}; x_{14}]$	3813–3814	70.5
...	...	...

Consider the following spline extrapolation for simulated self-similar traffic on the 3800–3850 ms segment using cubic B-spline (Fig. 6).

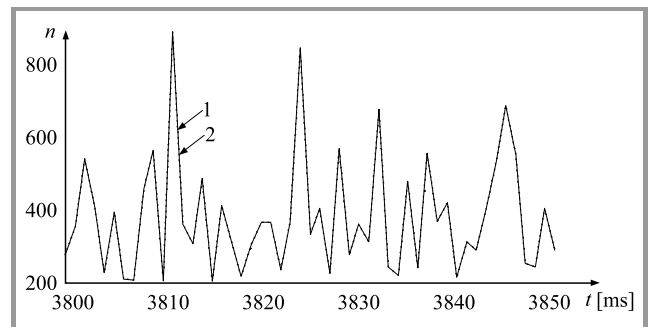


Fig. 6. Extrapolated self-similar traffic on the 3800–3850 ms segment using cubic B-spline functions: 1 – self-similar traffic, 2 – extrapolation of self-similar traffic using a cubic B-spline function.

Let us find the error values of the extrapolation of self-similar traffic using Eq. (21). The calculation results are shown in Table 3.

The use of cubic B-splines allows to reduce the error rate compared with the use of linear and cubic splines. According to the results of extrapolation of self-similar traffic on the 3800–3850 ms segment using the cubic B-splines shown in Fig. 6, the error of extrapolation is even negligible on the segment’s “bursts” of traffic intensity.

In general, the use of cubic B-splines allows to perform a short-term forecast of traffic parameters and to almost completely restore the “route” of self-similar traffic.

Let us compare the recovery of self-similar traffic using a linear, cubic, and cubic B-spline, as shown in Fig. 7. The research conducted shows that for the self-similar traffic considered, the use of cubic B-splines is most appropriate. The spline extrapolation method can be used to predict traffic characteristics in real time for short-term forecasts, and long-term forecasts that are based on a large amount of data.

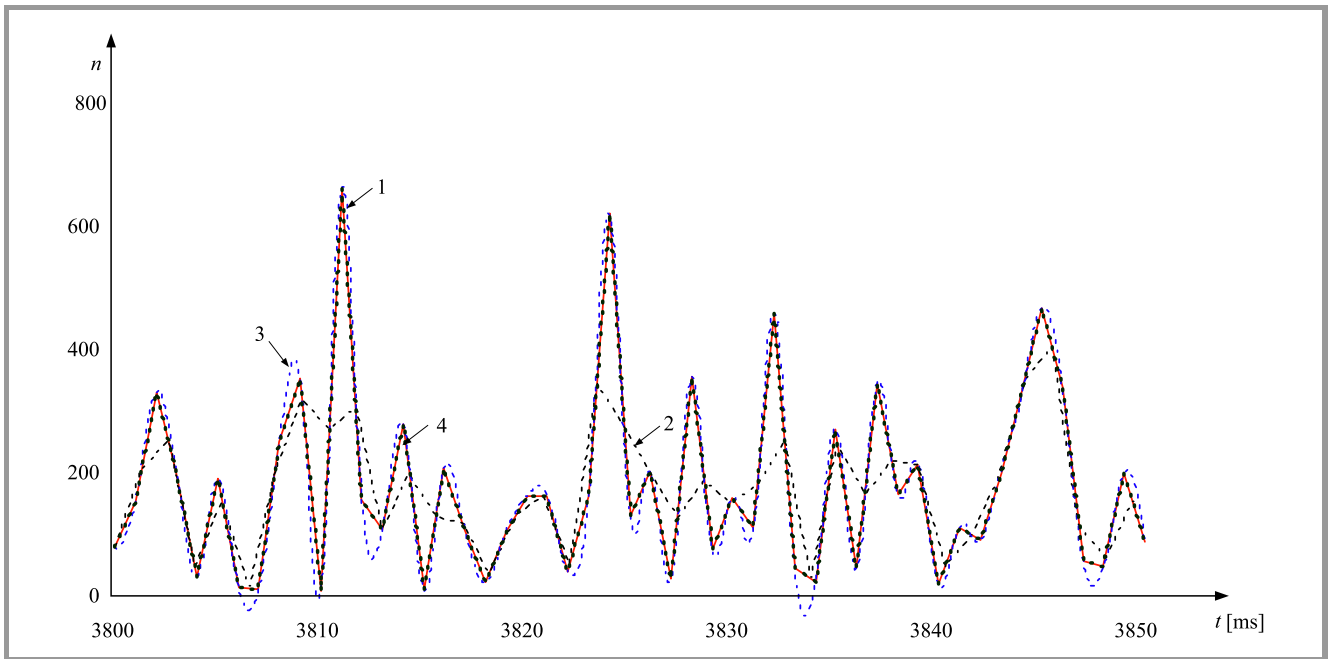


Fig. 7. Comparison of extrapolation results of self-similar traffic using linear, cubic and cubic B-spline on the 3800–3850 ms segment: 1 – self-similar traffic, 2 – extrapolation of self-similar traffic using linear spline function, 3 – extrapolation of self-similar traffic using cubic spline function, 4 – extrapolation of self-similar traffic using cubic B-spline function.

Table 3

The error of recovery of self-similar traffic based on cubic B-spline

Interval	Interval time [ms]	Error value
$[x_0; x_1]$	3800–3801	0.002
$[x_1; x_2]$	3801–3802	0.046
$[x_2; x_3]$	3802–3803	0.551
$[x_3; x_4]$	3803–3804	0.005
...	...	...
$[x_{10}; x_{11}]$	3810–3811	0.016
$[x_{11}; x_{12}]$	3811–3812	0.005
$[x_{12}; x_{13}]$	3812–3813	0.850
$[x_{13}; x_{14}]$	3813–3814	0.006
...	...	...

Short-term traffic forecast is usually associated with the ability to predict traffic characteristics in real time, in order to dynamically allocate network resources depending on traffic behavior, i.e. intensity bursts.

The results of such a forecasting approach allow to improve the network's QoS and to perform the optimal allocation of resources.

The long-term traffic forecast allows to obtain results that are based on a large amount of data, for example to choose the size of the buffer devices on the network nodes and to provide for the length of the packet queue in these devices when designing the network.

6. Conclusions

The proposed extrapolation method based on spline functions, has a number of advantages in comparison with the known methods. It is quite simple to implement, has a low error rate, and can also be used to control traffic in real time.

The practical significance of the results obtained is that the prediction of self-similar traffic will provide for the required amount of buffer devices, thereby avoiding overloads in the network and exceeding the standard QoS values.

Based on the results of traffic forecasting taking into account the maximum workloads of network nodes, practical recommendations can be given on traffic redistribution over IP networks, for example, the operation of TCP/IP protocol. Reduction of the delay, compared with TCP, allows to use a protocol without the guaranteed delivery of User Datagram Protocol (UDP). However, it is rather difficult to provide the required QoS using the UDP/TCP transport protocol only, since the reasons of the delays exist mostly at the network level [22]. Use of the proposed spline extrapolation method allows to perform a prediction of traffic characteristics, balance the load of network elements and improve the efficiency of network equipment.

Future research will need to focus on further improvement of traffic prediction accuracy, by developing the wavelet-extrapolation method based on wavelet functions.

References

- [1] 3GPP “Study on Scenarios and Requirements for Next Generation Access Technologies”, ETSI TR 38.913, V14.3.0, 2017 [Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138913/14.02.0060/tr_138913v140200p.pdf

- [2] 3GPP “Study on Architecture for Next Generation System”, TR 23.799 V14.0.0, 2016 [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3045>
- [3] V. V. Krylov and S. S. Samohvalova, *Teoriya teletrafika i ee prilozheniya (Teletraffic Theory and Its Applications)*. St. Petersburg: BHV-Petersburg, 2005, p. 288 (in Russian).
- [4] O. I. Sheluhin, A. V. Osin, and S. M. Smolski, *Samopodobie i Fraktaly. Telekomunikatsionnye Prilozheniya (Self-Similarity and Fractals. Telecommunication Applications)*. Moscow: Fizmatlit, 2008 (in Russian).
- [5] I. V. Strelkovskaya, I. N. Solovskaya, N. V. Severin, and S. A. Paskalenko, “Spline approximation-based restoration for self-similar traffic”, *Eastern-Eur. J. of Enterprise Technol.*, vol. 3/4 (87), pp. 45–50, 2017 (doi: 10.15587/1729-4061.2017.102999).
- [6] I. V. Strelkovskaya, I. N. Solovskaya, and N. V. Severin, “Modeling of self-similar traffic”, in *Proc. of 4th Int. Conf. on Appl. Innov. in IT ICAIIT-2016*, Koethen, Germany, 2016, vol. 4, no. 1, pp. 61–64 (doi: 10.13142/KT10004.23).
- [7] V. V. Popovsky and I. V. Strelkovskaya, “Accuracy of filtration procedures, extrapolation and interpolation of random processes”, *Problems of Telecommunications*, vol. 1, no. 3, pp. 3–10, 2011 [Online]. Available: http://pt.journal.kh.ua/2011/1/1/111_popovsky_estimation.pdf (in Russian).
- [8] I. Strelkovskaya, I. Solovskaya, and A. Makoganiuk, “Predicting characteristics of self-similar traffic”, in *Proc. of 3rd Int. Conf. on Inform. and Telecommun. Technol. and Radio Electronics UkrMiCo'2018*, Odessa, Ukraine, 2018.
- [9] J. V. Lambers and A. C. Sumner, *Explorations in Numerical Analysis*. World Scientific Publishing Company, 2018 (ISBN: 978-981-3209-96-1).
- [10] A. Messaoudi, R. Sadoka, and H. Sadok, “New algorithm for computing the Hermite interpolation polynomial”, *Numerical Algorithms*, vol. 77, no. 4 ppl. 1069–1092, 2018 (doi: 10.1007/s11075-017-0353-6).
- [11] I. Farago, A. Havasi, and Z. Zlatev, “Efficient implementation of stable Richardson Extrapolation algorithms”, *Comp. & Mathem. with Appl.*, vol. 60, no. 8, pp. 2309–2325, 2010 (doi: 10.1016/j.camwa.2010.08.025).
- [12] J. H. Ahlberg, E. N. Nilson, and J. I. Walsh, *The Theory of Splines and Their Applications*, 1st ed. Academic Press, 1967 (ISBN-13: 978-1483209524).
- [13] P. Sarigiannidis, K. Aprozidis, M. Louta, P. Angelidis, and T. Lagkas, “Predicting multimedia traffic in wireless networks: a performance evaluation of cognitive techniques”, in *Proc. 5th Int. Conf. on Inform., Intell., Sys. and Appl. IISA-2014*, Chania, Greece, 2014, pp. 341–346 (10.1109/IISA.2014.6878802).
- [14] V. Kumar and L. Vanajakshi, “Short-term traffic flow prediction using seasonal ARIMA model with limited input data”, *Eur. Transport Res. Review.*, vol. 7, no. 21, pp. 9–21 (doi: 10.1007/s12544-015-0170-8).
- [15] C. Li, Y. Han, Z. Sun, and Z. Wang, “A novel self-similar traffic prediction method based on wavelet transform for satellite Internet”, *EAI Endorsed Trans. on Ambient Syst.*, vol. 4, no. 14, pp. 1–7 (doi: 10.4108/eai.28-8-2017.153306).
- [16] T. H. H. Aldhyani and M. R. Joshi, “An integrated model for prediction of loading packets in network traffic”, in *Proc. 2nd Int. Conf. on Inform. and Commun. Technol. for Competitive Strateg. ICTCS'16*, Udaipur, India, 2016 (doi: 10.1145/2905055.2905236).
- [17] M. Oravec, M. Petras, and P. Pilka, “Video traffic prediction using neural networks”, *Acta Polytech. Hungarica*, vol. 5, no. 4, pp. 59–78, 2008 [Online]. Available: <https://www.uni-obuda.hu/journal/Oravec.Petras.Pilka.16.pdf>
- [18] F. C. Pereira, C. Antoniou, J. A. Fargas, and M. Ben-Akiva, “A meta-model for estimating error bounds in real-time traffic prediction systems”, *IEEE Trans. on Intell. Transport. Syst.*, vol. 15, no. 3, pp. 1310–1322 (doi: 10.1109/TITS.2014.2300103).
- [19] I. Klevecka, “Forecasting network traffic: a comparison of neural networks and linear models”, in *Abstracts of the 9th International Conference ‘Reliability and Statistics in Transportation and Communication’*, Latvia, Riga, 21–24 Oct., 2009. Riga: Transport and Telecommunication Institute, 2009, pp. 36–36 (ISBN: 978-9984-818-22-1).
- [20] Yu. S. Zavyalov, B. I. Kvasov, and V. L. Miroshnichenko, *Methods of Spline Functions*. Moscow: Nauka, 1980 (in Russian).
- [21] I. Strelkovskaya, “Application of cubic B-splines for synthesis of selective signals”, *Telecommun. and Radio Engin.*, vol. 66, no. 12, pp. 1047–1056, 2007 (doi: 10.1615/TelecomRadEng.v66.i12.10).
- [22] D. I. Comer, *Internetworking with TCP/IP*. Pearson Education Limited, 2013.



Irina Strelkovskaya holds a Ph.D. in Technical Sciences and is a Professor at the O. S. Popov Odessa National Academy of Telecommunications, Odessa, Ukraine. She is the author of more than 253 scientific and educational publications, including: 178 scientific articles and conference abstracts, 4 patents, 10 textbooks,

2 monographs, 59 educational and methodological manuals. Her research interests focus on analytic splines in areas with quasi-conformal boundaries, application of spline- and wavelet-apparatus in the telecommunications theory, tensor analysis to solve telecommunications-related problems.

 <https://orcid.org/0000-0003-2717-5014>

E-mail: i.strelkovskaya@onat.edu.ua

O. S. Popov Odessa National Academy of Telecommunications
Kuznechnaya, 1
65029 Odessa, Ukraine



Irina Solovskaya holds a Ph.D. degree and is an Associate Professor at the Department of Switching Systems at the O. S. Popov Odessa National Academy of Telecommunications, Odessa, Ukraine. She is the author of more than 70 scientific and educational publications, including 60 scientific articles and conference abstracts, 10 educational and methodological manuals and 1 monograph. Her research interests include the application of spline-apparatus and tensor method in solving telecommunications-related problems.

 <https://orcid.org/0000-0002-9904-5692>

E-mail: i.solovskaya@onat.edu.ua

O. S. Popov Odessa National Academy of Telecommunications
Kuznechnaya, 1
65029 Odessa, Ukraine



Anastasiya Makoganiuk is a teacher at the Department of Computer Science at the O. S. Popov Odessa National Academy of Telecommunications, Odessa, Ukraine. She is the author of more than 14 scientific articles and conference abstracts. Her research in-

terest includes the application of methods of spline-functions in the theory and practice of mobile communications networks.

 <https://orcid.org/0000-0002-7668-1653>

E-mail: a.makoganyuk@onat.edu.ua

O. S. Popov Odessa National Academy
of Telecommunications

Kuznechnaya, 1

65029 Odessa, Ukraine

Multifactor Authentication and Key Management Protocol for WSN-assisted IoT Communication

Tabassum Ara and M. Prabhakar

School of Computing and Information Technology, Reva University, Bangalore, India

<https://doi.org/10.26636/jtit.2019.134019>

Abstract—In this paper a novel multi-factor authentication protocol for IoT applications, relying on enhanced Rabin-assisted elliptic curve cryptography, biometric features and time stamping methods, is developed. Furthermore, a fuzzy verification algorithm has been developed to perform receiver-level user verification, making computation efficient in terms of computational overhead as well as latency. An NS2 simulation-based performance assessment has revealed that the multifactor authentication and key management models we have proposed are capable of not only avoiding security breaches, such as smart card loss (SCLA) and impersonation attacks, but can also ensure the provision of maximum possible QoS levels by offering higher packet delivery and minimum latency rates.

Keywords—multifactor authentication, IoT security, ECC, timestamp, one-way bio-hashing, fuzzy verifier, WSN.

1. Introduction

Advancement of new technologies has given rise, over the past few years, to a new paradigm known as the Internet of Things (IoT), relying on machine-to-machine (M2M) communications performed on a large scale. Functionally, these communication paradigms exploit Wireless Sensor Networks (WSN) with augmented routing protocols to simultaneously transmit data between multiple users or machines. Typically, IoT systems require secure data transmission or communication protocols across peers and, hence, a proper security system becomes an inevitable need [1], [2].

In IoT communication, the Internet Engineering Task Force (IETF) recommended certain protocols and standards to incorporate WSN into the Internet [3], such as IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN) [4] and Routing Over Low Power and Lossy Networks (ROLL) [5]. Functionally, IoT may comprise multiple sensor nodes connected to the Internet via a gateway. In such a scenario, the connected sensors would be accessible by any authorized entity, thereby ensuring remote access to and control over major applications. As a consequence, IoT communication remains vulnerable to attacks [2], [6]–[11]. Considering the criticality of data security across IoT ecosystems, the transmitted data must be

protected and secured across the sensor nodes and the entities connected to the network, by means of certain secure peer-to-peer channels. It is not feasible to apply Internet security measures directly to the IoT, due to WSN characteristics [12]. Nevertheless, efforts have been undertaken, such as IPsec [13] and IKEv2 [14], to ensure the security of IoT using Internet security models, with the resource-constrained nature of WSNs limiting their efficiency.

The classic uni-factor security models, such as password-based or bio-physical security systems are prone to being attacked, as the number of hacking techniques that exists is growing exponentially. Therefore, it becomes necessary to enable multi-factor assisted user (node) authentication to establish keys between nodes and the different authenticated stakeholders or entities. Furthermore, the existing efforts [15]–[17] have affirmatively stated that designing a secure IoT communication protocol is feasible. The objective may be achieved by enabling channel authentication and key management strategies requiring the remote entities to mutually authorize each other and to negotiate secret keys to assist the sensor nodes in avoiding active and passive attacks during the transmission [12]–[17]. Noticeably, even with certain security features deployed at the link layer of the IEEE 802.15.4 protocol stack, the openness of the Internet still causes significant vulnerability and, hence, demands certain robust key agreement and authentication schemes [12].

In this paper, a robust multi-factor authentication protocol has been developed that exploits efficacy of time-stamping, Rabin cryptosystem-assisted elliptic curve cryptography (ECC) and bio-hashing. To enable optimal user verification under real-time communication scenarios, a fuzzy-based verification model has been applied that learns the above mentioned features to provide access to a node for further communication. The proposed model has been implemented over a WSN with multiple cooperatively communicating or connecting nodes. NS-2 simulation has revealed that the proposed security model exhibits goods efficiency without incorporating any significant computational overheads or complexities.

The remaining sections of this paper are organized as follows. Section 2 presents the related work, which is fol-

lowed by the discussion of key existing approaches to IoT data security, given in Section 3. Section 4 shows the proposed system and its implementation. Section 5 presents the results and discussion, while conclusions concerning the overall results are given in Section 6.

2. Related Work

Realizing the shortcomings of single parameter-based security systems, Shin *et al.* [18] developed a two-factor authentication model where the prime focus was on employing an authenticated key agreement paradigm between users and IoT devices. However, this model was found to be limited due to its inability to address stolen or lost smart card attacks (SCLA), as well as offline password guessing and/or retrieval attacks relying on Brute Force, etc. To deal with these shortcomings, Wazid *et al.* [19] designed a new secure lightweight multi-factor remote user authentication scheme for hierarchical IoT networks (HloTN). They proposed a user-authenticated key management protocol (UAKMP). However, these approaches are complex. To reduce computational complexities, Sarvabhatla *et al.* [20] used a biometric feature authentication model for heterogeneous WSNs. However, personalized biometric traits preserved in a memory chip can be compromised due to smart card loss.

Challa *et al.* [21] developed a secure signature-based authenticated key establishment scheme for future IoT applications. Porambage *et al.* [22] developed a group key establishment protocol for secure multicast communications between resource-constrained sensor devices in IoT. However, the ambiguity of information shared between the interconnected nodes confines its efficiency in large networks. Ning *et al.* [23] developed a proof-based hierarchical authentication scheme for IoT. This model focused primarily on developing an U2IoT architecture (i.e. unit IoT and ubiquitous IoT), and eventually recommended an aggregated proof-based hierarchical authentication scheme (APHA). The authors combined directed path descriptors, homomorphism functions and Chebyshev chaotic maps for mutual authentication to ensure hierarchical access control.

Mick *et al.* [24] developed a lightweight authentication and secured routing (LASER) method for named data networks (NDN) used in smart city IoT applications. To enable computationally efficient routing, He *et al.* [25] developed an ECC-based RFID authentication scheme for the healthcare sector. Being an energy-constrained network, WSN-assisted IoT requires energy efficient and reliable transmission. Hence, Mohd *et al.* [26] developed lightweight block ciphers enhancing IoT security. The authors focused on both security and computational efficiency. Similar work was performed by Lu *et al.* [27], who developed a lightweight privacy-preserving data aggregation scheme, known as lightweight privacy-preserving data aggregation (LDPA). They relied on homomorphic Paillier encryption, Chinese remainder theorem and one-way hash

chain techniques to ensure efficient data gathering and to achieve a reduction in the false rate.

Jakalan *et al.* [28] designed a model called network security situation awareness (NSSA), where the focus was on assessing security situation-related elements and information originating from a multi-source heterogeneous networks. The authors considered four IoT security-related variables, such as context, attack, vulnerability, and network flow, which were then processed using the ontological concept to obtain the best security solution. In order to augment computational complexity, Diro *et al.* [29] recommended a lightweight cryptographic model – ECC. A similar effort was made by Yuan *et al.* [30], who developed a lightweight trust mechanism for IoT edge devices based on the fusion of multi-source feedback information. The use of multi-source feedback information for global trust estimation enabled avoiding a scenario in which a malicious node becomes a part of the network and, hence, enhance the security of the solution. This approach may be helpful in exploiting different node features to isolate the unauthorized node. However, this is done at the cost of increased computational overheads and latency.

Zahra *et al.* [31] employed Shibboleth, also known as the security and cross-domain access control protocol between fog a client and a fog node, to achieve secure communication between nodes, even under uncertain network conditions. Zheng *et al.* [32] utilized attribute-based encryption to enable data sharing over the network. In addition to the removal of attribute matching function, the use of the attribute bloom filter enabled hiding all attributes in the access control structure. However, the efficacy of their model could not be assessed in terms of node performance parameters. A similar effort was made by Want *et al.* [33]. Lei *et al.* [34] derived a closed-form model for assessing the probability of security-outage and its impact on throughput. They assessed their model to achieve better trade-offs between secure communication and energy-efficient data transmission over IoT. Lai *et al.* [35] derived a novel pairing-free data access control scheme by exploiting cipher text policy, attribute-based encryption (CP-ABE), where ECC was applied as encryption the method. Elhoseny *et al.* [36] designed a hybrid-cryptosystem using AES and RSA algorithms for diagnostic text data security in medical images. The authors applied 2-dimensional discrete wavelet transform and steganography concepts to ensure secure data transmission. In [37] Ruan *et al.* designed a leakage-resilient (LR) eCK security model for the password-based authenticated key exchange (PAKE) protocol. The authors applied the LR PAKE protocol with Diffie-Hellman key exchange LR storage (LRS).

3. Proposed Method

Unlike major existing methods, the proposed solution exploits multiple authorization schemes to design a robust security model, where each factor may function as a supplementary security layer, to ensure seamless communi-

cation. Considering the efficacy of a robust multi-factor authentication and key management system for IoT security, in this paper the emphasis is placed on developing an advanced security paradigm by amalgamating robustness of Rabin cryptosystem-assisted ECC, bio-hashing and the FLC- based verification mechanism.

3.1. Bio-Hashing

Over the past few years, the use of bio-traits as a security feature has increased significantly. Considering the increased threat of security breaches in the classic approaches, biometric features, such as fingerprint, retina and face detection, have been applied as additional security layers. The use of biometric traits as a supplementary security feature in conjunction with classic passwords or smart cards, has exhibited satisfactory results, making it one of the most sought-after techniques. Biometric features are deeply integrated and closely coupled with each individual. On the other hand, the irreplaceable nature of biometric features makes them most resistant to security breaches or to unauthenticated attempts to access specific systems or resources. Over the past few years, numerous research projects have been completed, relying on the efficiency of biometric data as security features [38], [39]. Biometric templates may be considered one of the pre-dominant privacy-preserving biometric approaches [39]. In practice, biometric template B and certain random constructs called secret key K are employed to generate a bio hash value (BV) $BV = \text{Bio2H}(K, B)$ during registration. Noticeably, here $\text{Bio2H}(K, B)$ defines a bilinear hashing of the biometric feature, such as a finger print, retina information, etc. $\text{Bio2H}(K, B)$ (two-factor bio hashing) pre-processes the biometric feature template B to preserve its integrity and accuracy. In other words, the pre-processing of B leaves the biometric feature concerned intact, even if certain minor variations in the input biometric signal are experienced. Once the BV has been generated, $\text{Bio2H}(K, B)$ compares the inner product of the random vector generated from the user specific secret key K and the feature vector extracted, against a predefined threshold value. During the verification phase, the same process as applied at the registration stage is followed. In this phase, the received biometric template or signal B' and the secret key specified by the user as a BV value $\text{Bio2H}(K, B')$ are estimated. Once the initial BV $\text{Bio2H}(K, B)$ and the re-estimated value of BV $\text{Bio2H}(K, B')$ have been obtained at the user end, they are compared to authenticate the user [39]. The key assumption for use of any bio-hashing technique is that the user has to be registered/on-boarded with at least one biometric feature captured (biometric template B), which is then converted into and stored in a suitable binary format (bits). In the proposed research, we have considered fingerprint bits $BioU_i$ to be biometric template B which, when combined with secret K at the registration end, is compared with received template B' combined with K as a signature to validate the user.

Table 1
Different cryptosystems and their features

Algorithm	Key exchange	Encryption/ decryption	Digital signature
Diffie Hellmen	Yes	Yes	No
DSA	No	No	Yes
RSA	Yes	Yes	Yes
ECC	Yes	Yes	Yes

Table 2
Notations used in this paper

Notation	Description
Bio2H	Two factor bio hashing
K	Secret key
B	Biometric feature template
B'	Received biometric feature template
P, Q	Points of elliptic curve
x, y	Coordinates on elliptic curve
M	Plain message
C	Cipher text
P	Prime number
GF	Galois field
a, b	Coefficients
G	Generator point
n, h	Cryptographic prime factor and cofactor
r	Random integer
P_r	Private key
P_u	Public key
U_{ID_i}	i -th user identification
PWU_i	i -th user password
l	Integer
U_i	i -th user
AD_j	Admin/sensor node
AD_{ID_j}	Admin's ID
AD_{xi}	Admin's x value
$h()$	Hash function
R_{shrd}	Shared random number
e_i, f_i, L_i	Hashed value
SCN_i	Smart card number
SCS	Smart card storing
Bio_{ui}	Bio information of user U_i
SC	Smart card
Z_i	Arbitrary number
EID	Estimated ID
GN	Gateway node
ID_{GN}	ID of gateway node
M_1, M_2, M_3, M_4	Intermediate values of messages
T	Delay threshold
MSG	Message
SK_{GN}	Session key of gateway node
DD_i, ZZ_i	Derived parameter
$e_i^{new}, f_i^{new}, g_i^{new}$	New estimated hash values

3.2. Rabin-Assisted Elliptic Curve Cryptography

There are numerous public key cryptosystems, such as RSA, DSA or Diffie Hellmen. The majority of these classic approaches suffer from huge computational overheads and time complexities (Table 1). Table 2 summarizes the notation used. Robustness of the ECC cryptosystem may be visualized based on the intricacy of the elliptic curve discrete log-arithm problem (ECDLP). Let the expression be $Q = kP$, where P and Q pertain to $F_p(a, b)$. With k and P values provided, it becomes easy to calculate Q . On the contrary, with P and Q pre-specified, it becomes more complicated to estimate k , especially when k is large. Noticeably, k states the discrete logarithm of Q to the base P , signifying the discrete logarithm problem for ECC. This computational complexity enhances attack resilience of ECC. The predominant process involved in ECC is known as point multiplication. Functionally, elliptic curve E is defined as:

$$y^2 = x^3 + ax + b. \quad (1)$$

In Eq. (1), the highest degree is 3. In the proposed method, ECC encrypts data M and generates ciphertext C and vice versa, using a certain finite set of points in the elliptic curve over $GF(p)$. Equation (1) (Weierstrass equation), $y^2 = x^3 + ax + b$ is used in conjunction with modulo p to generate points on the elliptic curve. Typically, the elliptic curve-specific variables are p, a, b, G, n, h, r where p states a prime number, a and b are the coefficients, and parameter G presents a generator point. The other parameters, such as n and h , define the cryptographic prime factor and the co-factor, respectively. Here, r is a random integer which is lower than n . The proposed model employs the finite field elliptic curve with modulo 263. A snippet of the applied method is given as Algorithm 1.

Algorithm 1: The proposed model concept

```

{
  Chose an elliptic curve having modulo  $p(y^2 = x^3 + ax + b \bmod p)$ 
  Assign the values of  $a$  and  $b$ , and of the coefficients.
  Estimate the value of  $y^2 = x^3 + ax + b \bmod p$ 
  {
    For  $x = 0$  to  $(p - 1)$ 
       $S = x^3 + ax + b \bmod p$ 
      For  $d = 0$  to  $(p - 1)/2$ 
        {
           $T = d^2 \bmod p$ 
          If  $T = S$ 
             $Y1 = d$  and  $Y2 = p - d$ 
          Else  $d = d + 1$ 
        }
       $x = x + 1$ 
    }
  }
   $(x, Y1), (x, Y2)$ 
}

```

In ECC, n is often multiplied by point generator G , eventually amounting to zero. Noticeably, generator point G may generate all points that may be potentially generated for

the defined elliptic curve. Next, a definite set of points is selected and the data are assigned to these points to be encrypted before the transmission. During the transmission, message M is at first encoded to point $P(M)$ from the finite set of points generated over the elliptic curve $Ep(a, b)$. To ensure high security levels, the selection of G is of paramount significance, where $G \in Ep(a, b)$. Typically, G and $Ep(a, b)$ are made public during the transmission. Both the transmitter and the receiver may choose private key Pr , with the help of which public key may be obtained as $P_u = P_r \times G$. In our implementation, random integer r is selected at first, which is then followed by the estimation of a cipher-text point using the receiver's public key. Thus, the cipher-text points obtained are $C = [(r.G), (M + r.Pur)]$. Now, the cipher-text is decrypted by multiplying the initial (first) point of the cipher-text pair $(r.G)$ by private key Prr , which is then subtracted from the second point of the cipher-text pair. The original message M is obtained as $M = (M + r.Pur) - Prr(r.G) = (M + r.PrrG) - Prr(r.G)$. A detailed presentation of the ECC algorithm may be found in [41]. Noticeably, unlike it is the case in classic cryptosystems, in this paper we have derived a hybrid method by combining the Rabin cryptosystem and ECC. Here, the Rabin cryptosystem has been used to generate the keys, while ECC has been employed to perform encryption-decryption of the message or the data being transmitted.

4. System Model

Unlike in the classic RSA, in the ECC-based cryptosystem presented in this paper we have designed a multifactor authentication and key management strategy. The proposed system incorporates the ECC cryptosystem, one-directional bio-hashing, time stamping and fuzzy verification. Such an approach results in a robust security solution enabling to avoid attacks, like SCLA and location tracking. The use of the ECC model is a lightweight crypto-solution. The use of the timestamp method enables mitigating session-specific temporary information attacks, eventually avoiding any tracking attacks. On the other hand, the use of the fuzzy verifier enables the model we have proposed to perform local verification of the users' passwords, thus offering a robust scheme for malicious node isolation. A snippet of the proposed multi-factor authentication and key management model is given below.

4.1. Model Description

As discussed above, the proposed system is initiated by relying on the Rabin-based ECC cryptosystem. In this preliminary phase, a sensor administration node ADM generates two prime numbers p and q , and estimates variable $N = pq$. Preserving (p, q) as the private key, it selects a master reference key x_{GN} in addition to an integer l ($2^4 \leq l \leq 2^8$) which is used at the receiver for local password verification, performed with the use of the fuzzy verifier. Estimating the receiver's master reference

key, hereinafter referred to as the public key, the cipher-text point is obtained as $C = [(r.G), (M + r.Pur)]$. The first point of the cipher-text pair $r.G$ is multiplied by private key Prr , and the result obtained is subtracted from the second point of the cipher-text pair.

Next, ADM selects an identity AD_{IDj} and estimates the secret key:

$$AD_{Xj} = h(AD_{IDj} \| x_{GN} \text{ for } AD_j (1 \leq j \leq M)).$$

Then, ADM generates a random number R_{shrd} , which is shared between the GN and the node AD_j . Node AD_j stores $AD_{IDj}AD_{Xj}$ and R_{shrd} in its memory.

4.1.1. User Registration and Login

During the user registration phase, user U_i performs the following steps to get registered with ADM.

User U_i transmits the selected identity U_{IDi} along with their personal credentials to ADM, by means of a specific secure channel.

Upon receiving information from U_i , ADM verifies whether U_{IDi} exists in the table or in the database. If the verification is positive, ADM acknowledges U_i to select an updated or a new identity; else it generates an arbitrary number x_i and then estimates variable $d_i = h(U_{IDi} \| x_{GN} \| x_i)$ and $L_i = h(SC_{N_i} \| x_{GN})$. Here, SC_{N_i} defines the SC number belonging to a user or a node. Estimating the values of d_i and L_i , ADM enables SCS.

U_i connects the smart card into a card reader and then enters U_{IDi} , PWU_i and imports $BioU_i$, SC selects a random number r_i and estimates the attributes:

$$BioZ_i = Bio2H(r_i, BioU_i),$$

$$e_i = h(h(U_{IDi} \| PWU_i \| BioZ_i) \bmod l),$$

$$f_i = d_i \oplus h(U_{IDi} \| PWU_i \| BioZ_i) \text{ and}$$

$$g_i = L_i \oplus h(U_{IDi} \oplus PWU_i \oplus BioZ_i).$$

Upon performing user registration, the user requires a login to the system. This step occurs when user U_i intends to access sensor data.

First, U_i connects SC and feeds its unique identity U_{IDi}^* , along with password-related information PWU_i^* . In addition, it requires that fingerprint information $BioU_i$, be embedding or attached. Now, the deployed SC estimates $BioZ_i^* = Bio2H(r_i^*, BioU_i)$ in addition to $e_i^* = h(h(U_{IDi}^* \| PWU_i^* \| BioZ_i^*) \bmod l)$. This process ensures that in the case of $e_i^* \neq e_i$, the card will reject user U_i , without granting access to sensor data.

Next, the SC generates an arbitrary number z_i along with a timestamp T_{stamp1} , and calculates:

$$d_i^* = f_i \oplus h(U_{IDi}^* \| PWU_i^* \| BioZ_i^*),$$

$$L_i^* = g_i \oplus h(U_{IDi}^* \oplus PWU_i^* \oplus BioZ_i^*),$$

$$M_1 = (U_{IDi} \| SC_{N_i} \| z_i)^2 \bmod n,$$

$$M_2 = h(d_i^* \| L_i^* \| z_i \| T_{stamp1}).$$

Finally, U_i selects the ID of the j -th sensor AD_{IDj} which it intends to access, and the SC estimates value $EID_j =$

$AD_{IDj} \oplus h(U_{IDi} \| z_i \| T_{stamp1})$, which is then appended and transmitted to the GN as $MSG_1 = \langle M_1, M_2, T_1, EID_j \rangle$.

4.1.2. Authentication

To perform a secure or valid mutual authentication, and to agree on the session key, the proposed model performs the following processes.

Upon receiving MSG_1 from U_i , GN at first decrypts M_1 using p and q to obtain U'_{IDi} , SC'_{N_i} , and z'_i . x_i is retrieved as per U'_{IDi} which is then validated in the entries with SC'_{N_i} . If these values fail to match, GN rejects the request and terminates the process. Otherwise, it estimates:

$$L'_i = h(SC'_{N_i} \| x_{GN},$$

$$d'_i = h(U'_{IDi} \| x_{GN} \| x_i),$$

$$z'_i = M_2 \oplus h(d'_i \| T_{stamp1}),$$

$$M'_2 = h(d'_i \| L'_i \| z'_i \| T_{stamp1}).$$

In the case of $M'_2 \neq M_2$, GN aborts the current session. On the contrary, for $M'_2 = M_2$ it estimates:

$$AD'_{IDj} = EID_j \oplus h(U_{IDi} \| z_1 \| T_{stamp1}),$$

$$AD(X_j)' = h(AD'_{IDj} \| x_{GN}),$$

$$M_3 = h(U'_{IDi} \| AD'_{IDj} \| ID_{GN} \| AD'_{Xj} \| z'_i \| T_{stamp2}),$$

$$M_4 = U'_{IDi} \oplus h(ID_{GN} \| AD'_{Xj} \| T_{stamp2}),$$

$$M_5 = z_i \oplus h(U'_{IDi} \| AD'_{IDj} \| AD'_{Xj} \| T_{stamp2}).$$

Next, GN transmits message $MSG_2 = \langle ID_{GN}, M_3, M_4, T_{stamp1} \rangle$ to AD_j .

In the next step, AD_j verifies whether $|T_{stamp3} - T_{stamp1}| \leq \Delta T$ exists, where T_{stamp3} signifies the current timestamp. If the verification result is positive, AD_j stops the session, otherwise it estimates:

$$U_{IDi}^{**} = M_4 \oplus h(ID_{GN} \| AD_{Xj} \| T_{stamp2}),$$

$$z_i^{**} = M_5 \oplus h(U_{IDi}^{**} \| AD_{IDj} \| AD_{Xj} \| T_{stamp2}),$$

$$M'_3 = h(U_{IDi}^{**} \| AD_{IDj} \| ID_{GN} \| AD_{Xj} \| z_i^{**} \| T_{stamp2}).$$

Now, AD_j aborts the connection in the case of $M'_3 \neq M_3$. Otherwise, it confirms the authenticity of U_i and GN. Next, AD_j estimates the following:

$$SK_j = h(U_{IDi}^{**} \| AD_{IDj} \| z_i^{**} \| z_j),$$

$$M_6 = h(SK_j \| AD_{Xj} \| z_j \| T_{stamp3}),$$

$M_7 = z_i^{**} \oplus z_j$, where z_j is a random number generated by AD_j .

In the next phase, AD_j forwards $MSG_3 = \langle M_6, M_7, T_3 \rangle$ to GN.

In the following phase, GN verifies whether $|T_{stamp4} - T_{stamp3}| \leq \Delta T$. If not, GN terminates the session. Otherwise, it estimates:

$$z'_j = M_7 \oplus z'_i,$$

$$SK_{GN} = h(U_{IDi} \| AD_{IDj} \| z'_i \| z'_j),$$

$$M'_6 = h(SK_{GN} \| AD'_{Xj} \| z'_i \| T_{stamp3}).$$

In the case of $M'_6 \neq M_6$, GN terminates the connection and estimates $M_8 = h(SK_{GN} \| U'_{IDi} \| d'_i \| z'_i)$. Thus, the gateway node GN transmits $MSG_4 = \langle M_7, M_8 \rangle$ to U_i .

Once receiving MSG_4 , U_i estimates:

$$z_j^* = M_7 \oplus z_i,$$

$$SK_i = h(U_{IDi} \| AD_{IDj} \| z_i \| z_j^*), \text{ and}$$

$$M_8' = h(SK_i \| U_{IDi} \| d_i \| z_j^*).$$

If $M_8' \neq M_8$, U_i terminates the session. Otherwise, it assumes that gateway node GN and AD_j are authentic. At this instant, a common session key $SK_i = SK_j = SK_{GN}$ is established between the participating U_i , GN and AD_j .

5. Results and Discussion

To examine the effectiveness of the proposed model, reference works [38], [39], where three factor-based authentication was proposed, have been taken into consideration. The overall results have been assessed in terms of qualitative and quantitative, or empirical outcomes. As far as the qualitative assessment is concerned, the proposed method has been examined for its ability to avoid different attack scenarios. On the other hand, in the quantitative assessment, the model has been integrated with a WSN routing protocol, developed and simulated using Network Simulator-2, version (NS-2). The outcomes have been examined in terms of latency, packet delivery and packet loss rates.

5.1. Resistance to SCLA Attacks

Consider that attacker A obtains SC information containing $\langle e_i, f_i, g_i, SC_{N_i}, l, n, r_i, Bio2H(.,.), Hash1() \rangle$ of valid user U_i .

Attacker A may be able to guess user the credentials of U_{IDi}^* as well as PWU_i^* , which may help in estimating $e_i^* = h(h(U_{IDi}^* \| PWU_i^* \| BioZ_i) \bmod l)$. However, they cannot retrieve the correct value of U_{IDi}^* and PWU_i^* due to e_i being a “fuzzy verifier”. This novelty enables the proposed system to avoid an SCLA attack. In addition, the proposed model is capable of withstanding an SCLA Type II attack as well.

Consider that the attacker node A has identified the message $MSG_1 = \langle M_1, M_2, T_1, EID_j \rangle$ transmitted by U_i when logging in, where:

$$d_i^* = f_i \oplus h(U_{IDi}^* \| PWU_i^* \| BioZ_i^*),$$

$$L_i^* = g_i \oplus h(U_{IDi}^* \oplus PWU_i^* \oplus BioZ_i^*).$$

In the above expression, g_i is obtained from SC of U_i SC. The complexity of the quadratic residue problem makes it infeasible for the attacker to estimate R_1 from the value $M_1 = (U_{IDi} \| SC_{N_i} z_i)^2 \bmod n$, and hence it prevents A from estimating value $M_2^* = h(d_i^* \| L_i^* \| z_i \| T_1)$, that is required to verify the user with its U_{IDi}^* and PWU_i^* . This approach makes the proposed model resilient to SCLA Type II attacks.

5.1.1. Resistance KSSTIA and User Impersonation Attacks

In the existing methods, their authors applied a static value $h(U_{IDi}' \| AD_{Xj}')$ to secure the ephemeral arbitrary num-

bers, where variable AD_{Xj}' signifies the node's key shared between AD_j and GW. Consequently, the revealing of ephemeral random number z_i compromises the static value $h(U_{IDi}' \| AD_{Xj}')$, which eventually leads to revealing or compromising the ephemeral arbitrary numbers in other key management or authentication sessions. To alleviate the risk, a multi-factor authentication and key management policy with timestamp and one-way hashing concepts is incorporated into the proposed model. such an approach allows the proposed model to avoid KSSTIA type attacks.

In user impersonation attack scenarios, attacker A is prevented from performing any user impersonation attacks. Consider that attacker A manages to retrieve U_i 's SC card and extracts information like $\langle e_i, f_i, g_i, SC_{N_i}, l, n, r_i, Bio2H(.,.), Hash1() \rangle$. Furthermore, let's assume that A has already identified or obtained the messages communicated in the previous authentication sessions. In such cases, our proposed security model forces A to have all authorizing factors including PWU_i , SC, and biometric information for generating a certain valid message $MSG_1 = \langle M_1, M_2, T_1, EID_j \rangle$. Practically, the key required to facilitate the authenticity of the user U_i refers to value $M_2 = h(d_i^* \| L_i^* \| z_i \| T_1)$. The vital constructs of M_2 encompass $d_i^* = f_i \oplus h(U_{IDi}^* \| PWU_i^* \| BioZ_i^*)$ and $L_i^* = g_i \oplus h(U_{IDi}^* \oplus PWU_i^* \oplus BioZ_i^*)$. However, without providing the password for that user PWU_i , as well as SC and biometric information, the intruder or attacker A cannot estimate d_i^* or L_i^* .

5.1.2. Resistance to Gateway Impersonation Attacks

In the proposed system, attacker A is capable of impersonating the user by pretending GW to be either user U_i or AD_j . To impersonate GW as AD_j , the intruder A must estimate $M_3 = h(U_{IDi}' \| AD_{IDj}' \| ID_{GN} \| AD_{Xj}' \| z_i' \| T_2)$. However, without having precise information about $h(U_{IDi}' \| x_{GN})$, it is impossible for the intruder to estimate the value of M_3 . On the other hand, the use of the hash algorithm and the time stamping technique means that intruder A will not be able to retrieve any significant information from messages obtained from the previous authentication sessions. On the contrary, impersonating as GN to user U_i , the intruder requires estimating a valid factor $M_8 = h(Z_{Session_{GN}} \| U_{IDi}' \| d_i' \| z_j')$. To achieve it, the intruder requires having information about z_i that further helps in the estimation of $Z_{Session_{GN}} = h(U_{IDi}' \| AD_{IDj}' \| z_i \| z_j)$. To retrieve the value of z_i , A requires knowing the secret key of GN. This condition cannot be fulfilled, as the secret key is preserved or protected by the administrator. An in-depth analysis shows that A can impersonate by decrypting $M_1 = (U_{IDi} \| SC_{N_i} z_i)^2 \bmod n$. However, this is a highly tedious and challenging task due to the computational complexity of Rabin-assisted ECC and its allied quadratic residue problem (QRP). Therefore, the proposed security model or the allied protocol eliminates any possibility of a gateway node impersonation attack.

5.1.3. Resistance to Modification and Replay Attacks

In the proposed method, intruder A is unable to make any modification to messages $MSG_1 = \langle M_1, M_2, T_1, EID_j \rangle$, $MSG_2 = \langle ID_{GN}, M_3, M_4, M_5, T_2 \rangle$, $MSG_3 = \langle M_6, M_7, T_3 \rangle$, or $MSG_4 = \langle M_7, M_8 \rangle$. Consider that attacker A is capable of intercepting any one of the message chunks. Then, it may be able to modify the same and transmit them further. However, in the proposed method, each message is protected by means of a hash value that is estimated using a secret value. Therefore, the attacker cannot retrieve the message. For illustration, in MSG_1 , attacker A cannot estimate the value of $M_2 = h(d_i^* || L_i^* || z_i || T_1)$, because $d_i^* = f_i \oplus h(U_{IDi}^* || PWU_i^* || BioZ_i^*)$ and $L_i^* = g_i \oplus h(U_{IDi}^* \oplus PWU_i^* \oplus BioZ_i^*)$ are secret values which cannot be estimated without knowing PWU_i , SC or the biometric feature. In the case of any modification, the receiver can detect it when checking the correctness of the hash value in each message. Thus, our proposed system can be described as being modification resilient.

In IoT, mobile nodes may be used. In such a scenario, the attacker may try to replay a stale message transmitted by a certain user. In the proposed method, the timestamp feature allows to resist any replay attacks.

5.1.4. Resistance to Insider Attacks and Verification of Credentials

Typically, nodes or users may sign up to various applications or information systems using similar passwords. Should an insider somehow get access to the password, they may use it for impersonating the user and getting access to their data. User U_i submits U_{IDi} when signing up

or registering. Therefore, the insider will not be able to achieve the user's password.

While attacking the authentication server, verification information, e.g. the password, may be retrieved or stolen. The proposed method enables the server to retain such attributes as $\langle U_{IDi}; SC(N_i); x_i; \text{personal credential} \rangle$ and does not store any password-related information. Therefore, even after gaining access to the authentication server, the attacker cannot obtain the user's password-related information.

5.1.5. Mutual Verification, Session Key Attack, and Anonymity

In WSN-based IoT, mutual authentication is required between the nodes. In practice, the intruder is not capable of easily retrieving $M_2 = h(d_i^* || L_i^* || z_i || T_1)$ without having the genuine private key of the user, d_i^* and L_i^* . In such a case, the gateway node GN is not able to authenticate U_i by verifying the precision of M_2 . In the same way, user U_i can verify gateway node GN by verifying the correctness of $M_8 = h(Z_{Session_{GN}} || U'_{IDi} || d'_i || z'_j)$. In this scenario, user U_i and the GN can authenticate each other. Furthermore, AD_j authorizes GN by checking the correctness of $M_3 = h(U'_{IDi} || AD'_{IDj} || ID_{GN} || AD'_{Xj} || z'_j || T_2)$. Similarly, gateway node can authorize AD_j by verifying the correctness of $M_6 = h(Z_{Session_j} || AD_{Xj} || z_j || T_3)$. The proposed security model offers seamless mutual authentication between GN and AD_j .

During mutual authentication, session key $Z_{Session} = h(U_{IDi} || AD_{IDj} || z_i || z_j)$ is formed between user U_{IDi} and AD_j to secure future communication. Noticeably, the security of $Z_{Session}$ relies on the secrecy of the random numbers involved. In fact, these values remain protected by the

Table 3
Comparison of the proposed solution and other methods

Security features	Amin <i>et al.</i> [39]	Yeh <i>et al.</i> [42]	Hue <i>et al.</i> [43]	Jiang <i>et al.</i> [16]	Gope <i>et al.</i> [44]	Proposed
User untraceability	○	○	○	●	○	●
Replay attack		○		●	●	●
User impersonation attack	●	○	○	○	●	●
Gateway node impersonation attack	○	○	○	○	●	●
Sensor node impersonation attack	●	●	●	●	●	●
De-synchronization attack	○	○	○	○	●	●
Support of dynamic node addition	●	○	○	○	●	●
Insider attack	●		○	○	●	●
Stolen smartcard attack	●	○	○	○	●	●
User anonymity	●	○	○	●	●	●
Identity guessing attack	●	○	○	●	●	●
Support of three-factor security	●	○	○	○	○	●
Supports correct password update	●	○	○	○	○	●
Session key disclosure attack	○	●	●		○	●
Bio hashing	○	○	○	○	○	●
Quality of service	○	○	○	○	○	●
Rabin-assisted ECC	○	○	○	○	○	●
Fuzzy verification	○	○	○	○	○	●

secret values shared between WSN nodes participating in the exchange of each message. Let the session key be $Z_{Session} = h(U_{IDi} || AD_{IDj} || z_i || z_j)$ which is somehow known to intruder A. Even though A knows the session key, they cannot estimate any future or past session keys using $Z_{Session}$ as the session key itself is secured with the help of one-way hash function $Hash1()$. In addition, the random number used $\langle z_i, z_j \rangle$ may be different in each session, and, therefore, the proposed model offers resistance to any session key attacks.

A situation may be experienced when intruder A retrieves messages transmitted between the users and tries to identify the user. Now, observing our proposed model, it can be found that it integrates user identity in $M_1 = (U_{IDi} || SC_{N_i} || z'_i)^2 \bmod n$. To retrieve user identity U_{IDi} , the intruder requires the knowledge of the secret key when executing Rabin-assisted ECC of gateway GW. In practice, it is impossible to retrieve, as it is already stored by administrator. On the other hand, the computational-complexity of the quadratic residue problem (QRP) makes it near impossible for the intruder to obtain U_{IDi} by decrypting the value of $M_1 = (U_{IDi} || SC_{N_i} || z'_i)^2 \bmod n$. Table 3 compares the proposed solution with other similar methods described in the referenced publications.

5.2. Simulation Results

To assess the efficiency of the proposed (secure) WSN routing protocol, we simulated it with a distributed sensor network comprising 50 nodes cooperating across the

region. Furthermore, each node was assigned a radio range of 200 m. To examine the effectiveness of the proposed secure routing protocol in avoiding malicious or attack nodes, two compromising or attacker nodes were incorporated in the simulation environment (Table 4).

Figure 1 presents the packet delivery ratio (PDR) performance of our proposed security model and a routing protocol without any security features. As depicted, performance of the proposed model was nearly equivalent to that of the model without any security. This signifies that the proposed approach is lightweight and may be well suited for any WSNs. Similar performance-related results are visualized in Fig. 2, showing that the proposed security model exhibits a much lower packet loss rate than the classic routing protocol.

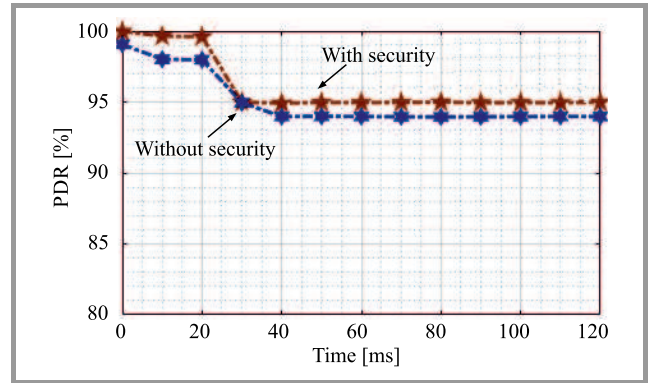


Fig. 1. Trade-off analysis for packet delivery ratio.

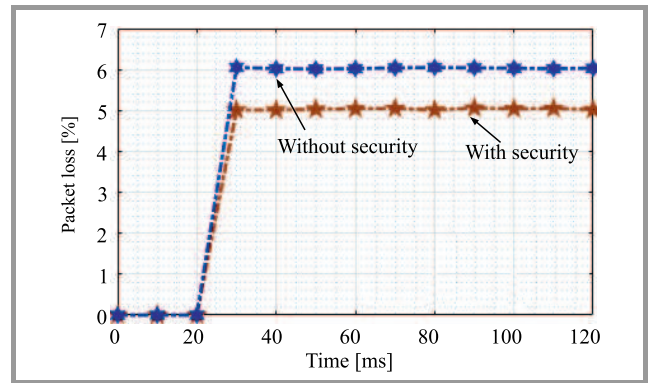


Fig. 2. Trade-off analysis for packet loss.

Figure 3 presents the rate of transmission over varying simulation times or transmission periods. Here, one may observe that the transmission rate of the proposed security method is nearly the same as that of the classical or native routing protocol. This shows that incorporation of the proposed multi-factor authentication and key management approaches does not affect the transmission rate significantly. In WSN-based IoT systems, ensuring timely data transmission or delivery is of utmost significance. On the other hand, it is hypothesized that the inclusion of any security models may impose computational overheads, causing increased latency. In such cases, assessment of the efficiency

Table 4
Nodes characteristics in simulation environment

Parameter	Value
MAC	IEEE 802.15.4
PHY	802.15.4PHY
Antenna	Omni-directional
Radio range	200
Sensor nodes	50
Efficiency of RF power amplifier	0.47
Link margin	40 dB
Gain factor	30 dB
Power density of AWGN channel	-134 dBm/Hz
Noise figure (receiver)	10 dB
Path loss	3-5 dB
Carrier frequency	2.5 GHz
Bandwidth	20 kHz
BER performance	$1 \cdot 10^{-3}$
Transmitter circuit power consumption	98.2 mW
Receiver circuit power consumption	112.6 mW
Antenna gain of transceiver	5 dB
Routing table update period per round	5
Routing table size	100
Transmission rate	2-10 kb/s
Packet size	2 kbits
Transmission probability of each node	0.8

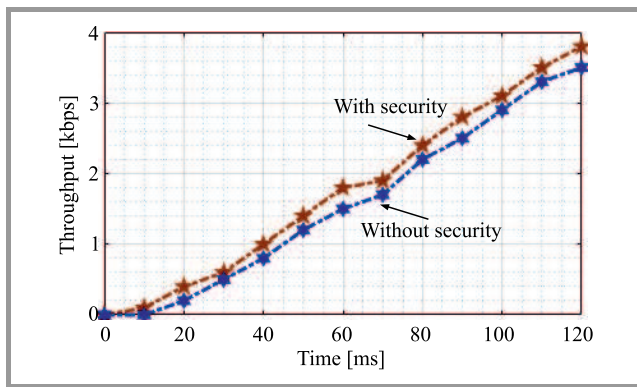


Fig. 3. Trade-off analysis for throughput.

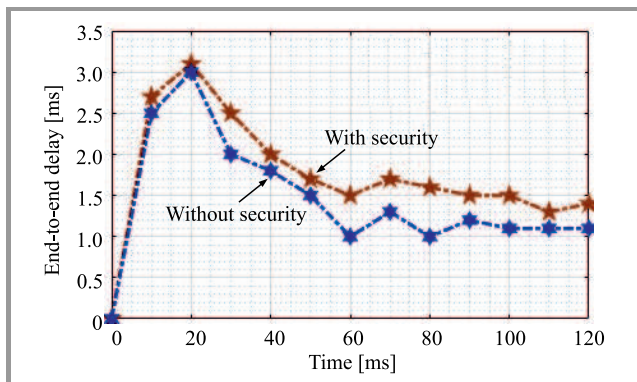


Fig. 4. Trade-off analysis for end-to-end delay.

of the proposed secure protocol in terms of delay, is a must. Figure 4 presents the end-to-end delay performance of the proposed secure routing protocol.

6. Conclusions

The proposed protocol is characterized by a delay performance that is similar to that of classic or native WSN routing protocols. The proposed secure routing protocol is capable of ensuring the QoS required without suffering any significant delays, computational overheads, packet losses, retransmission rates or multifactor authentication. The key management model allows not only to avoid such security breaches as SCLA and impersonation attacks, but is also capable of ensuring the maximum possible QoS levels by guaranteeing higher packet delivery and minimum latency rates. The proposed model has exhibited satisfactory performance for WSN-based IoT systems.

References

- [1] S. Hong *et al.*, "SNAIL: An IP-based wireless sensor network approach to the Internet of Things", *IEEE Wirel. Commun.*, vol. 17, no. 6, pp. 34–42, 2010 (doi: 10.1109/WMC.2010.5675776).
- [2] J. Granjal, E. Monteiro, and J. S. Silva, "Security in the integration of low power Wireless Sensor Networks with the Internet: A survey", *Ad Hoc Netw.*, vol. 24, pp. 264–287, 2015 (doi: 10.1016/j.adhoc.2014.08.001).
- [3] Z. Sheng *et al.*, "A survey on the IETF protocol suite for the Internet of Things: Standards, challenges, and opportunities", *IEEE Wirel. Commun.*, vol. 20, no. 6, pp. 91–98, 2013 (doi: 10.1109/WMC.2013.6704479).
- [4] "6LoWPAN Status Pages", 6LoWPAN Working Group [Online]. Available: <http://tools.ietf.org/wg/6lowpan/> (accessed on Jan. 15, 2017).
- [5] "Roll Status Pages", ROLL Working Group [Online]. Available: <http://tools.ietf.org/wg/roll/> (accessed on Oct. 18, 2018).
- [6] R. Roman and J. Lopez, "Integrating wireless sensor networks and the Internet: A security analysis", *Internet Res.*, vol. 19, no. 2, pp. 246–259, 2009 (doi: 10.1108/10662240910952373).
- [7] J. Astorga, E. Jacob, N. Toledo, and J. Unzilla, "Enhancing secure access to sensor data with user privacy support", *Comput. Netw.*, vol. 64, pp. 159–179, 2014 (doi: 10.1016/j.comnet.2014.02.002).
- [8] J. Qi, X. Hu, Y. Ma, and Y. Sun, "A hybrid security and compressive sensing-based sensor data gathering scheme", *IEEE Access*, vol. 3, pp. 718–724, 2015 (doi: 10.1109/ACCESS.2015.2439034).
- [9] Z. Fu, X. Sun, Q. Liu, L. Zhou, and J. Shu, "Achieving efficient cloud search services: Multi-keyword ranked search over encrypted cloud data supporting parallel computing", *IEICE Trans. Commun.*, vol. E98-B, no. 1, pp. 190–200, 2015 (doi: 10.1587/transcom.E98.B.190).
- [10] Z. Fu, F. Huang, X. Sun, A. V. Vasilakos, and C.-N. Yang, "Enabling semantic search based on conceptual graphs over encrypted outsourced data", *IEEE Trans. Serv. Comput.*, pp. 1–1, 2016 (doi: 10.1109/TSC.2016.2622697).
- [11] H. Li, D. Liu, Y. Dai, and T. H. Luan, "Engineering searchable encryption of mobile cloud networks: When QoE meets QoP", *IEEE Wirel. Commun.*, vol. 22, no. 4, pp. 74–80, 2015 (doi: 10.1109/WMC.2015.7224730).
- [12] R. Roman, C. Alcaraz, J. Lopez, and N. Sklavos, "Key management systems for sensor networks in the context of the Internet of Things", *Comp. & Elect. Engin.*, vol. 37, no. 2, pp. 147–159, 2011 (doi: 10.1016/j.compeleceng.2011.01.009).
- [13] S. Raza *et al.*, "Securing communication in 6LoWPAN with compressed IPsec", in *Proc. 7th Int. Conf. on Distrib. Comput. in Sensor Syst. & Workshops DCOSS 2011*, Barcelona, Spain, 2011, pp. 1–8 (doi: 10.1109/DCOSS.2011.5982177).
- [14] S. Ray and G. Biswas, "Establishment of ECC-based initial secrecy usable for IKE implementation", in *Proc. World Congr. on Engin. WCE 2012*, London, U.K., 2012, vol. 1, pp. 530–535 (ISBN: 978-988-19251-3-8).
- [15] S. Kumari, M. K. Khan, and M. Atiquzzaman, "User authentication schemes for wireless sensor networks: A review", *Ad Hoc Netw.*, vol. 27, pp. 159–194, 2015 (doi: 10.1016/j.adhoc.2014.11.018).
- [16] Q. Jiang, J. Ma, X. Lu, and Y. Tian, "An efficient two-factor user authentication scheme with unlinkability for wireless sensor networks", *Peer-Peer Network. Appl.*, vol. 8, no. 6, pp. 1070–1081, 2014 (doi: 10.1007/s12083-014-0285-z).
- [17] D. He, N. Kumar, and N. Chilamkurti, "A secure temporal-credential based mutual authentication and key agreement scheme with pseudo identity for wireless sensor networks", *Inform. Sci.*, vol. 321, pp. 263–277, 2015 (doi: 10.1016/j.ins.2015.02.010).
- [18] S. Shin and T. Kwon, "Two-factor authenticated key agreement supporting unlinkability in 5G-integrated Wireless Sensor Networks", *IEEE Access*, vol. 6, pp. 11229–11241, 2018 (doi: 10.1109/ACCESS.2018.2796539).
- [19] M. Wazid, A. K. Das, V. Odelu, N. Kumar, M. Conti, and M. Jo, "Design of secure user authenticated key management protocol for generic IoT networks", *IEEE Internet of Things J.*, vol. 5, no. 1, pp. 269–282, 2018 (doi: 10.1109/JIOT.2017.2780232).
- [20] M. Sarvabhatla and C. S. Vorugunti, "A secure biometric-based user authentication scheme for heterogeneous WSN", in *Proc. 4th Int. Conf. of Emerg. Appl. of Inform. Technol.*, Kolkata, India, 2014, pp. 367–372 (doi: 10.1109/EAIT.2014.23).
- [21] S. Challa *et al.*, "Secure signature-based authenticated key establishment scheme for future IoT applications", *IEEE Access*, vol. 5, pp. 3028–3043, 2017 (doi: 10.1109/ACCESS.2017.2676119).

- [22] P. Porambage, C. Schmitt, P. Kumar, A. Gurtov, and M. Ylianttila, "Two-phase authentication protocol for wireless sensor networks in distributed IoT applications", in *Proc. IEEE Wirel. Commun. and Networ. Conf. WCNC 2014*, Istanbul, Turkey, 2014, pp. 2728–2733 (doi: 10.1109/WCNC.2014.6952860).
- [23] Z. Mahmood, H. Ning, and A. Ghafoor, "Lightweight two-level session key management for end user authentication in Internet of Things", in *Proc. IEEE Int. Conf. on Internet of Things (iThings) and IEEE Green Comput. and Commun. (GreenCom) and IEEE Cyber, Phys. and Social Comput. and IEEE Smart Data (SmartData)*, Chengdu, China, 2014, pp. 323–327, 2016 (doi: 10.1109/iThings-GreenCom-CPSCoM-SmartData.2016.78).
- [24] T. Mick, R. Tourani, and S. Misra, "LASER: Lightweight authentication and secured routing for NDN IoT in smart cities", *IEEE Internet of Things J.*, vol. 5, no. 2, pp. 755–764, 2018 (doi: 10.1109/JIOT.2017.2725238).
- [25] D. He, N. Kumar, N. Chilamkurti, and J.-H. Lee, "Lightweight ECC based RFID authentication integrated with an ID verifier transfer protocol", *J. of Medical Syst.*, vol. 38, no. 10 (doi: 10.1007/s10916-014-0116-z).
- [26] B. J. Mohd and T. Hayajneh, "Lightweight block ciphers for IoT: energy optimization and survivability techniques", *IEEE Access*, vol. 6, pp. 35966–35978, 2018 (doi: 10.1109/ACCESS.2018.2848586).
- [27] R. Lu, K. Heung, A. H. Lashkari, and A. A. Ghorbani, "A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT", *IEEE Access*, vol. 5, pp. 3302–3312, 2017 (doi: 10.1109/ACCESS.2017.2677520).
- [28] A. Jakalan, "Network security situational awareness", *The Int. J. of Comp. Sci. and Commun. Secur. (IJCSCS)*, vol. 3, pp. 61–67, 2013 [Online]. Available: https://www.researchgate.net/publication/256932621_Network_Security_Situational_Awareness
- [29] A. A. Diro, N. Chilamkurti, and N. Kumar, "Lightweight cybersecurity schemes using elliptic curve cryptography in publish-subscribe fog computing", *Mobile Netw. and Appl.*, vol. 22, no. 5, pp. 848–858, 2017 (doi: 10.1007/s11036-017-0851-8).
- [30] J. Yuan and X. Li, "A reliable and lightweight trust computing mechanism for IoT edge devices based on multi-source feedback information fusion", *IEEE Access*, vol. 6, pp. 23626–23638, 2018 (doi: 10.1109/ACCESS.2018.2831898).
- [31] S. Zahra *et al.*, "Fog computing over IoT: A Secure deployment and formal verification", *IEEE Access*, vol. 5, pp. 27132–27144, 2017 (doi: 10.1109/ACCESS.2017.2766180).
- [32] H. Zheng, J. Wu, B. Wang, and J. Chen, "Modified ciphertext-policy attribute-based encryption scheme with efficient revocation for PHR system", *Mathem. Problems in Engin.*, vol. 2017, article ID 6808190, pp. 1–10 (doi: 10.1155/2017/6808190).
- [33] S. Wang, D. Zhao, and Y. Zhang, "Searchable attribute-based encryption scheme with attribute revocation in cloud storage", *PLoS ONE*, vol. 12, no. 8, 2017 (doi: 10.1371/journal.pone.0183459).
- [34] H. Lei *et al.*, "Performance analysis of physical layer security over generalized- K fading channels using a mixture gamma distribution", in *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 408–411, 2016 (doi: 10.1109/LCOMM.2015.2504580).
- [35] J. Lai, R. Deng, C. Guan, and J. Weng, "Attribute-based encryption with verifiable outsourced decryption", *IEEE Trans. Inform. Forensics and Secur.*, vol. 8, no. 8, pp. 1343–1354, 2013 (doi: 10.1109/TIFS.2013.2271848).
- [36] M. Elhoseny *et al.*, "Secure medical data transmission model for IoT-based healthcare systems", *IEEE Access*, vol. 6, pp. 20596–20608, 2018 (doi: 10.1109/ACCESS.2018.2817615).
- [37] O. Ruan, J. Chen, and M. Zhang, "Provably leakage-resilient password-based authenticated key exchange in the standard model", *IEEE Access*, vol. 5, pp. 26832–26841, 2017 (doi: 10.1109/ACCESS.2017.2776160).
- [38] R. Amin and G. P. Biswas, "A secure light weight scheme for user authentication and key agreement in multi-gateway based wireless sensor networks", *Ad Hoc Netw.*, vol. 36, pp. 58–80, 2016 (doi: 10.1016/j.adhoc.2015.05.020).
- [39] R. Amin *et al.*, "Design of an anonymity-preserving three-factor authenticated key exchange protocol for wireless sensor networks", *Comput. Netw.*, vol. 101, no. C, pp. 42–62, 2016 (doi: 10.1016/j.comnet.2016.01.006).
- [40] I. Natgunanathan, A. Mehmood, Y. Xiang, G. Beliaikov, and J. Yearwood, "Protection of privacy in biometric data", *IEEE Access*, vol. 4, pp. 880–892, 2016 (doi: 10.1109/ACCESS.2016.2535120).
- [41] T. D. Pramila-Bai, S. A. Rabara, and A. V. Jerald, "Elliptic curve cryptography based security framework for Internet of Things and cloud computing", *Int. J. of Comp. Sci. & Technol.*, vol. 6, no. 3, pp. 223–229, 2015 (ISSN: 2229-4333) [Online]. Available: https://www.researchgate.net/publication/305913586_Elliptic_Curve_Cryptography_based_Security_Framework_for_Internet_of_Things_and_Cloud_Computing
- [42] H. L. Yeh *et al.*, "A secured authentication protocol for wireless sensor networks using elliptic curves cryptography", *Sensors*, vol. 11, no. 5, pp. 4767–4779, 2011 (doi: 10.3390/s110504767).
- [43] K. Xue, C. Ma, P. Hong, and R. Ding, "A temporal-credential-based mutual authentication and key agreement scheme for wireless sensor networks", *J. of Netw. and Comp. Appl.*, vol. 36, no. 1, pp. 316–323, 2013 (doi: 10.1016/j.jnca.2012.05.010).
- [44] P. Gope and T. Hwang, "A realistic lightweight anonymous authentication protocol for securing real-time application data access in wireless sensor networks", *IEEE Trans. on Indust. Electron.*, vol. 63, no. 11, pp. 7124–7132, 2016 (doi: 10.1109/TIE.2016.2585081).



Tabassum Ara is a Research Scholar in Computer Science at Reva University, Bangalore, Karnataka, India, with her research interests focusing on IoT, security and WSN. With 18 years of academic experience, she holds B.Eng., M.Sc. and M.Tech. degrees and is an Assistant Professor at the Department of Computer Science

at HKBK College of Engineering in Bangalore.

 <https://orcid.org/0000-0002-3378-3713>

E-mail: tabuara@gmail.com

School of Computing and Information Technology
Reva University
Bangalore, India



M. Prabhakar received his M.Sc. and Ph.D. degrees in Computer Engineering from Anna University, Chennai. He boasts 21 years of teaching experience and is currently working as an Associate Professor at the School of Computing & Information Technology, REVA University, Bangalore, India. His areas of research interest include adhoc networks and cybersecurity.

E-mail: prabhakar.m@reva.edu.in

School of Computing and Information Technology
Reva University
Bangalore, India

A Novel Graph-modification Technique for User Privacy-preserving on Social Networks

Seyedeh Hamideh Erfani and Reza Mortazavi

School of Engineering, Damghan University, Damghan, Iran

<https://doi.org/10.26636/jtit.2019.134319>

Abstract—The growing popularity of social networks and the increasing need for publishing related data mean that protection of privacy becomes an important and challenging problem in social networks. This paper describes the (k, l) -anonymity model used for social network graph anonymization. The method is based on edge addition and is utility-aware, i.e. it is designed to generate a graph that is similar to the original one. Different strategies are evaluated to this end and the results are compared based on common utility metrics. The outputs confirm that the naïve idea of adding some random or even minimum number of possible edges does not always produce useful anonymized social network graphs, thus creating some interesting alternatives for graph anonymization techniques.

Keywords—graph-modification, social networks, privacy-preserving publication of data, graph anonymization, database security.

1. Introduction

In recent years, the extensive use of social networks, such as Facebook, Twitter or MySpace, in family or friendship communications, as well as in political, social, economic, educational, cultural and religious activities, has led many researchers to focus on various aspects of this highly utilized social communication tool (social network). Online social networks have provided many data sharing platforms. Due to the quickly increasing popularity of social network sites on the Web, disclosure of user identity becomes an important problem. Since the risk has a devastating impact on the daily life of the people involved (for example, disclosure of sensitive data, such as e-mails, instant messages, or private relationships), protecting user privacy has become an important and challenging problem in online social networks. Anonymization of the social network structure [1], [2] is a common approach to protecting user privacy.

A social network may be modeled as a graph, where vertices represent individuals, organizations or users, and edges represent connections, relationships between users or information flows [3]. Many researchers have focused on studying the problem of user privacy in online social networks [4]–[8]. To improve the security of social networks,

various social network data anonymization techniques have been proposed [9]–[15]. These anonymization approaches are designed based on the principle of k -anonymity.

Anonymization methods are used in undirected social network graphs. Furthermore, these approaches anonymize a social network by inserting and/or deleting edges and vertices in the graph. Since the social network graph structure has been changed by anonymization methods, the utility-related value of social network decreases. For example, if the relationship between two users A and B is removed, user A cannot retrieve sensitive data of user B, and user B will not be able to share sensitive data with user A. If user A is removed, their existence is ignored in the social network. Therefore, developing anonymization methods to protect user privacy in social networks without experiencing the loss of information continues to be an outstanding problem that is still difficult to solve [16].

Ideally, an anonymized social network should protect the privacy of individuals with a minimum loss in its utility-related value, ensuring that an analysis based on anonymized social network data is very similar to that based on its original counterpart [17]. Therefore, the problem of maintaining the utility of data is very important in the process of anonymization of social network data. By referring high levels of utility of data, we mean strict preservation of the pure information bits carried by the original social network data. The information that is distorted and differs, significantly, from the original will probably provide unreliable analysis results. However, most of the existing anonymization techniques fail to generate anonymized social network data with a high degree of utility [17]. Thus, it is important to understand and model the utility of social network data being published by relying on utility-aware metrics [18].

Social networks are usually anonymized based on some computational privacy models. This paper applies the (k, l) -anonymity model which was initially defined by Feder *et al.* in 2008 [19]. The authors suggested to use edge addition to implement the model. However, this initial definition suffers from some practical issues that were later addressed in the improved version created by Stokes and Torra [20]. In this definition, a graph is called (k, l) -anonymous if for

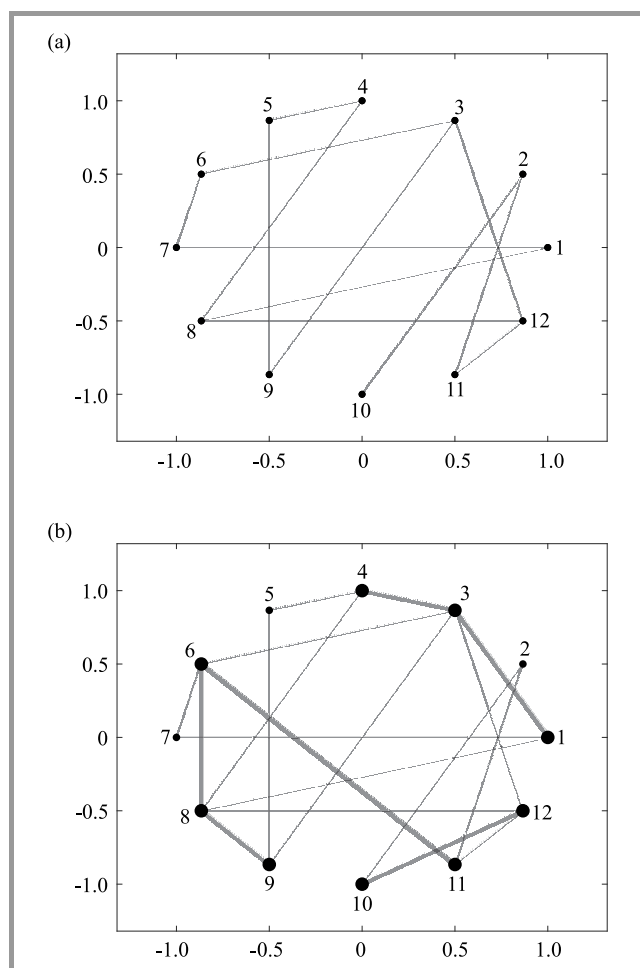


Fig. 1. The original graph (a) and its anonymized version based on edge addition with highlighted new edges (b).

every vertex v in the graph there exist at least $k - 1$ other vertices that share at least l of their neighbors with v . The authors suggest to introduce some fake nodes or even to remove all risky nodes to produce an anonymous graph. Although the modification procedures are easy to realize, the usefulness of these suggestions is not evaluated by using utility measures commonly referred to in the graph anonymization literature. Unfortunately, there is the risk that node removal deletes all nodes of the graph, since removing some risky nodes (along with their connected edges) introduces new risky nodes. Additionally, introducing new nodes does not satisfy the privacy requirement in practice, since they do not correspond to distinct individuals [20]. Usually, the amount of changes made to the social graph is used to control the utility distortion in structural-based anonymization [21]–[24]. However, this metric does not consider the impacts on the social links' structure, which has a serious impact on the graph properties [25]. For instance, consider the graph shown in Fig. 1a. If an attacker knows exactly two neighbors of a victim node in the graph, his chance to identify the victim distinctly in the graph is, on average, more than 90%

in average. This means that in order to remedy the risk based on naive node deletion, the algorithm has to remove almost all nodes in the original graph. However, adding some new edges, as shown in the Fig. 1b, deprives the attacker of this possibility.

In this paper, a mathematical model is applied to enforce the privacy requirement of (k, l) -anonymity to minimize the cost function related to the distortion of utility of the anonymization process. The aim of this study is to propose a general mathematical model to realize (k, l) -anonymity in social network graphs. To the best of our knowledge, this is the first general procedure that addresses the (k, l) -anonymity problem in graphs.

In summary, we offer the following main contributions:

1. We develop a general mathematical programming model to produce a (k, l) -anonymous graph for $l \geq 1$ and $k \geq 2$.
2. We linearize the anonymization constraints in our mathematical programming approach, which makes it solvable by mixed integer programming (MIP) solvers.
3. We show that the naïve idea of adding some random or even the minimum number of possible edges does not always produce useful anonymized social network graphs.
4. We compare the performance of the proposed approach with the existing approach. The results show that our approach changes slightly the key characteristics of the graph and produces more useful graphs.

The remainder of this paper is organized as follows. Section 2 reviews some related works. Section 3 defines the preliminary concepts that our work is based upon. The proposed method is described in Section 4 and is then evaluated in Section 5. Finally, Section 6 concludes the paper.

2. Related Work

Here, a number of graph-modification techniques are reviewed as graph anonymization methods. To preserve the privacy of users in the process of publishing data through anonymization methods, the process can be fully random or may be subject to some constraints, meaning that these methods are called random perturbation methods and constrained perturbation methods, respectively.

Random perturbation methods are based on introducing random noise to the original data. In the case of graphs, two main approaches exist: (a) rand add/del that randomly adds and deletes edges from the original graph, and (b) rand switch that exchanges edges between pairs of nodes. Random perturbation techniques are generally the simplest form of graph modification techniques and are less complex. Thus, they are able to deal with large networks, although

they do not offer privacy guarantees, but a probabilistic re-identification model [26] only. For example, Hay *et al.* [27] proposed a method, known as random perturbation, to anonymize unlabeled graphs using the rand add/del strategy that randomly removes p edges and then randomly adds p fake edges. Stokes and Torra [28] professed that a suitable selection of the eigenvalues in the spectral method can perturbate the graph while keeping its most significative edges. The authors in [29] presented a strategy which aimed to preserve the most important edges in the network, trying to maximize data utility while achieving the desired privacy level.

Several methods have been presented in the category of constrained perturbation methods, such as k -anonymity and extended k -anonymity. These methods provide privacy guarantees, but the level of privacy they offer may strongly depend on the knowledge of the adversary. The most basic knowledge of adversary is based on vertex degree [26]. The k -anonymity model was introduced in [30] and [31] for privacy preservation of structured or relational data. The k -anonymity model indicates that the attacker cannot distinguish between different k records, although he manages to find a group of quasi-identifiers. Therefore, the attacker can not re-identify an individual with a probability greater than $1/k$ [26]. There are some other models developed attempting to extend the k -anonymity model to overcome some particular disadvantages. Feder *et al.* [19] called a graph (k, l) -anonymous if, for every vertex in the graph, there exist at least k other vertices that share at least l of their neighbors. Severe weaknesses of that work were addressed by Stokes and Torra [20].

3. Preliminaries

In this section the preliminary concepts upon which this work is built will be introduced, including: graph anonymization and utility assessment of an anonymous graph.

3.1. Graph Anonymization

Similarly to other works, we model the social network as a simple graph, in which the degree of a vertex represents the number of its neighbors. Let $G = (V, E)$ be a social network graph, where V denotes the set of vertices and E represents the set of edges. As in [19], we assume, throughout this paper, that the social network graph is undirected, unweighted, and contains no self-loops, because this is an important category of graphs to study. Most of the social networks, such as Facebook, LinkedIn, Orkut and so on, allow only bidirectional links and are thus instances of such simple graphs.

Consider a simple graph and an attacker who knows that a target person and some number of their friends form a community. In the released graph, the attacker could find such a community to narrow down the set of nodes that might correspond to the target individual. The goal

of the anonymization method is to prevent the attacker from identifying individuals based on their immediate neighbors. To achieve this, we use the (k, l) -anonymity definition introduced by Feder *et al.* in 2008. In this definition, it is required that for every node in the graph, some subset of its neighbors should be shared by other nodes. In this way, an attacker who knows some subset of the neighbors of the target individual and is even capable of pinpointing them in the graph, will not be able to distinguish the target individual from other nodes in the network that share this subset of neighbors [19]. More formally the (k, l) -anonymity property is defined as follows.

Definition 1 (k, l) -anonymity [19]. A graph $G = (V, E)$ is (k, l) -anonymous if for each vertex $v \in V$, there exists a set of vertices $U \subset V$ not containing v such that $|U| \geq k$ and for each $u \in U$ the vertices u and v share at least l neighbors.

3.2. Utility Assessment of an Anonymous Graph

We now discuss how to measure the utility loss of an anonymized graph based on topological features of the graph. For a given original graph $G(V, E)$, $\hat{G}(V, E)$ is the (k, l) -anonymous version, such that the utility of the anonymized graph should be as close as the utility of the original graph [25]. Zhou and Pei [23], [32] consider the total number of added edges and the number of vertices that are not in the neighborhood of the target vertex and are linked to the anonymized neighborhood for the purpose of anonymization. Another work in [21] controls the utility loss by preferring the largest frequent subgraphs during anonymization to impose less graph modification. The total cost of anonymization is still calculated by the amount of changes made during perturbation [25]. Although, based on this criterion, the structural properties of the social network graph are ignored, the number of changes to minimize the information loss of network graph is still important.

When a social network graph is anonymized, the resulting graph loses some utility compared with the original graph [33]. Some attempts have been made to preserve the structural properties of a social network graph during anonymization. Research done by [18], [34] considers local community structure as the subject of utility preservation. In [34], the authors proposed an approach in which the graph is partitioned by a local structure.

In this paper, we have used the average vertex degree (AVD), the average path length (APL), and the average closeness centrality (ACC) of the anonymized graph in comparison with the original graph, in order to evaluate the proposed method. These three indices, along with the degree distribution, are considered to be standard measures in graph-analysis studies. These criteria are defined as follows (Table 1).

Consider an unweighted and undirected graph G having the set of vertices V that $|V| = n$. Assume $d(v_1, v_2)$, where $v_1, v_2 \in V$ denote the shortest distance between v_1 and v_2 . The degree of vertex v is denoted as $\deg(v)$.

Table 1
Notations used in the paper

Notation	Description
n	Number of graph vertices
k, l	Anonymization parameters
$\deg_i$	Degree of v_i
AVD	Average vertex degree
APL	Average path length
ACC	Average closeness centrality
CV	Candidate vertices
MIP	Mixed integer programming
PSPL	Pairwise shortest path length

Average vertex degree. The average vertex degree is defined as the average degree of all network nodes [35]. Then, the average vertex degree AVD_G is:

$$AVD_G = \frac{1}{n} \cdot \sum_{v \in V} \deg(v).$$

Average path length. The average path length is defined as the average number of steps along the shortest paths between all pairs of reachable network nodes. The average path length APL_G is defined as:

$$APL_G = \frac{1}{n \cdot (n-1)} \cdot \sum_{i \neq j} d(v_i, v_j).$$

Average closeness centrality. The closeness centrality of a node is a measure of centrality in a network, calculated as the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the graph. The average closeness centrality in a graph is given by the average closeness centrality of all nodes.

4. The proposed Method

In this section, we describe how to satisfy (k, l) -anonymity in a given graph $G = (V, E)$, dividing the process into two main parts. First, solving the mathematical model that the anonymization problem is defined by, using a mixed integer programming in which the best edges are found to be added to G to produce the anonymous graph $\hat{G}$. Second, post-processing that removes unnecessary edges from $\hat{G}$. In Algorithm 1, the two main parts of the proposed method are shown. The function accepts a simple graph $G(V, E)$, anonymization parameters k and l as inputs and produces an anonymous graph $\hat{G}(V, E')$. The number of vertices is stored in n in line 1. The pairwise shortest path length matrix corresponding to graph G is computed in line 2. In line 3, for each newly added edge e_{ij} that connects v_i to v_j , its associated cost related to its addition is computed. The anonymous graph is produced in lines 4–10. For the case of $l = 1$ (line 4), the optimal solution is obtained by using the MIP mathematical model in line 5 that is introduced in Subsection 4.1. For the case of $l > 1$, before solving MIP problem in line 9 that yields G_2 , a set of candidate vertices (CV) is computed using the *candidateVertices* function in line 8, which is shown in Algorithm 2. In line 10, all

Algorithm 1: The pseudo-code of the proposed method

Input: $G(V, E)$: original graph, k, l : anonymization parameters
Output: $\hat{G}(V, E')$: anonymized graph

```

1  $n = |V|$ 
2  $PSPL \leftarrow$  pairwise shortest path length matrix of graph  $G$ 
3 compute  $costs(i, j), \forall \{v_i, v_j\} \in E'$ 
4 if  $l == 1$  then
5    $\hat{G} \leftarrow MIPsolver(G, PSPL, k, l, n, costs, \deg)$ 
6 else
7   if  $l > 1$  then
8      $CV \leftarrow candidateVertices(G, PSPL, k, l)$ 
      // Algorithm 2
9      $G_2 \leftarrow MIPsolver(G, PSPL, k, l, n, costs, \deg, CV)$ 
10     $\hat{G} \leftarrow postProcess(G, G_2, k, l)$ 
      // Algorithm 3
11 return  $\hat{G}$ 

```

unnecessary added edges are removed from G_2 using the *postProcess* function (Algorithm 3), then its output is saved in $\hat{G}$, which is returned in line 11.

4.1. Solving Mathematical Model

This paper uses edge addition as the graph-modification technique to produce the anonymous graph. In order to produce a useful graph, a general mathematical model of the problem is introduced that takes into account the k minimum number of different vertices sharing l of their neighbors. The model uses the following components:

Definition of sets: The indexes of vertices $v \in V$ in the graph are saved in S .

Fixed parameters and constants: Three parameters n, k , and l represent the number of vertices, and k and l of the anonymization model, respectively. Parameter c_{ij} is the cost of adding the edge e_{ij} that connects v_i to v_j . The cost matrix $C = [c_{ij}]$ is symmetric, i.e., $c_{ij} = c_{ji}$. It is equal to 0 for connected vertices in the original graph, and assumes a positive value for not connected ones. Therefore, to reduce the amount of data passed to the solver, only the upper triangular part of C is applied in practice. Additionally, $\deg_i$ denotes the degree of v_i in the original graph.

Independent problem variables: The solution consists of connected vertices. The binary decision variable x_{ij} determines the connectivity of v_i and v_j in the produced anonymized graph, where $x_{ij} = 1$, if and only if the related vertices are to be connected. In order to decrease the space complexity of the final model, we only consider x_{ij} for $j > i, i, j \in S$, since the graph is undirected ($x_{ij} = x_{ji}$) and loop free ($x_{ii} = 0$).

Objective function: The objective function is to minimize the aggregate cost of change with respect to the original graph, i.e.:

$$\min_{x_{ij}} \sum_{i,j \in S, i < j} c_{ij} x_{ij}.$$

Constraints: The constraints fall into two categories: original edge preserving constraints, and anonymization constraints.

1. Original edge preserving constraints: none of the existent edges in the original graph cannot be removed, i.e.:

$$x_{ij} = 1, \quad \forall e_{ij} \in E.$$

In order to speed up the computation, the warm-start strategy is used in which $x_{ij} = 1$ for all of the connected vertices v_i and v_j .

Anonymization constraints: assume that $N_l(j)$ is the l neighbors of v_j , i.e. $N_l(j) = \{S_j | \#S_j = l, \forall i \in S_j, (v_i, v_j) \in E\}$. These constraints enforce that each l neighbors of a vertex are also the neighbors of at least k different vertices in the anonymized graph:

$$\sum_{w \in S' \setminus S'_j} \left(\prod_{i \in S'_j} x_{iw} \right) \geq k \quad \forall S'_j \in N_l(j), j \in S.$$

This function is non-linear due to the multiplication of x -variables. It is required to replace each nonlinear term $\prod_{i \in S'_j} x_{iw}$ with a new variable $z(i_1, i_2, \dots, w)$

where $i_n \in S'_j$ and add linearization constraints in the following manner:

$$\begin{aligned} 0 &\leq z(i_1, i_2, \dots, w) \leq x_{i_1 w} \\ 0 &\leq z(i_1, i_2, \dots, w) \leq x_{i_2 w} \\ &\vdots \\ 0 &\leq z(i_1, i_2, \dots, w) \leq x_{i_{|S'_j|} w} \\ z(i_1, i_2, \dots, w) &\geq x_{i_1 w} + x_{i_2 w} + \dots + x_{i_{|S'_j|} w} - (|S'_j| - 1) \end{aligned}$$

In the case of $l = 1$, as we introduced in [36], for each k the constraints can be simplified to the constraints that impose the minimum degree of each vertex to be at least k , i.e.:

$$\sum_{i < j} x_{ij} + \sum_{j < i} x_{ji} \geq k \quad \forall i \in S, \deg_i < k.$$

It is also notable that for a fixed w , $\prod_{i \in S'_j} x_{iw} = \prod_{i \in S''_j} x_{iw}$

where S''_j is defined in the following:

$$S''_j = S'_j - \{i | i \text{ is connected to } w\}.$$

2. Run-time constraint: according to Definition 1, we call the graph $G = (V, E)$, (k, l) -anonymous, if for each vertex $v_i \in V$, there exist at least l vertices, which are simultaneously connected to v_i and also

to at least $k - 1$ other vertices. In other words, if $v_j \in V$ connects to p vertices and also q other vertices connect to those p vertices, so that $|p| \leq l$ and $q < k$, then the graph G is not (k, l) -anonymous. In this situation, the MIP solver for finding the minimum cost (k, l) -anonymous graph needs to connect the vertex v_j to at least $l - p$ number of best vertices, so that all of the l vertices are connected to k other vertices. The MIP solver selects the best vertices (vertices with the minimum cost of adding their corresponding edges to the original graph) from all vertices in the graph. Therefore, for large and dense graphs and in the case of $l > 1$, implementation of the proposed method will take a lot of time. In this paper, in order to solve the problem in a reasonable time, we have decided that for all combination of l vertices of the graph G , we propose a set of vertices to the MIP solver that is called candidate vertices. These CV corresponding to l vertices are obtained from the union of their adjacent vertices. The CV is obtained by using Algorithm 2 (line 8 of Algorithm 1). The *CandidateVertices* function gets the original graph G , the PSPL matrix corresponding to graph G and anonymization parameters k and l . The CV is generated as the output of the function. In Algorithm 2, for all combination of l vertices of the graph G , a sorted vector D is produced in line 2 in which $\text{PSPL}(v_{i_k}, :)$ is the i_k -th row in the pairwise shortest path length matrix of G . Since the main diagonal of the PSPL matrix is zero, the first l elements in sorted vector D will be zero. In line 3, *validIdx* is the index of vertices obtained from $(l + 1)$ -th component of D and for threshold number ($Th \geq k$) of it. That Th is a positive number that determines the number of candidate vertices for all combinations of l vertices. *validIdx* introduces vertices that are considered for connecting to vertices $v_{i_1}, v_{i_2}, \dots, v_{i_l}$. In line 4, the CV is obtained as a matrix with $l + 1$ dimensions that $\text{CV}(i_1, i_2, \dots, i_l, \text{validIdx}) = 1$. For example, for $l = 2$ and $Th = k$, D is obtained by Eq. (1) that the CV would be a $n \times n \times n$ matrix that for each 2-combination (v_1, v_2) , $v_1, v_2 \in V$, the vertex v is a candidate vertex for connecting to both v_1 and v_2 provided by $\text{CV}(v_1, v_2, v) = 1$, for $v \in \text{validIdx}$. Therefore, CV proposes candidate vertices to the MIP solver and the MIP solver finds the best vertices from v to connect them to the vertices $\{v_{i_1}, v_{i_2}, \dots, v_{i_l}\} \in V$.

$$D = \text{sortIndex}(\text{PSPL}(v_{i_1}, :) \cdot \text{PSPL}(v_{i_2}, :)), \forall i_1, i_2 \in S.$$

$$\text{validIdx} = D[3 : 3 + k].$$

$$\text{CV}(i_1, i_2, \text{validIdx}) = 1.$$

(1)

In this study, two different variants of the general mathematical model have been applied to the original graph. The difference between these variants consists in the different cost functions that are to be minimized in the objective function. Specifically, the following models are considered:

Table 2
Structural properties of synthetic and real graphs

Dataset	Vertices	Edges	AVD	APL	ACC
SF50	50	96	3.8400	2.8433	0.0073
SF100	100	196	3.9200	2.9048	0.0035
SF200	200	396	3.4449	3.9600	0.0015
RA82	82	94	2.2927	5.7350	0.0022
RA129	129	178	2.7597	4.8307	0.0017
RA176	176	238	2.7045	5.4292	0.0011
karate	34	78	4.5882	2.4082	0.0129
dwt_72	72	75	3.0833	8.2254	0.0018
lesmis	77	254	6.5974	2.6411	0.0051
can_96	96	336	8.0000	4.4105	0.0024
polbooks	105	441	8.4000	3.0788	0.0032
football	115	613	10.6609	2.5082	0.0035

roduced in Subsection 3.2. The weaker the modification of these measures caused by the method concerned, the better the anonymized graph is suited for future investigation.

5.1. Datasets

The proposed method is applied to a number of synthetic and real graph datasets that are available online to see its performance in different topologies. Structural properties of these graphs are shown in Table 2. For each graph, the number of vertices, the number of edges and other struc-

tural properties described in Subsection 3.2, such as AVD, APL, and ACC, are reported. The synthetic datasets are summarized as follows:

- **SF** – a scale-free dataset based on the Barabasi's model. This graph is a connected graph, where vertex degrees are drawn from a power-law distribution similar to real-world social networks;
- **RA** – a random network based on the Erdos-Renyi model in which vertices are connected based on probability p .

Additionally, six real datasets are tested in a similar manner to assess our method in different topologies:

- karate – a social network of friendships between 34 members of a karate club at a US university in the 1970s;
- dwt – this collection consists of thirty matrix patterns collected by Gordon Everstine of the David W. Taylor Naval Ship Research and Development Center, Bethesda, MD, USA. These patterns were collected from various US military and NASA users of NASA's structural engineering package NASTRAN, for use as a benchmark collection for variable bandwidth re-ordering heuristics;
- lesmis – co-appearance weighted network of characters in the novel "Les Miserables";
- can – a graph that has a symmetric pattern made of cans;

Table 3

The amount of AVD error for different values of k and l and for the proposed method and for the Feder *et al.* [19] model applied to synthetic graphs. In each experiment, the best value is highlighted

Graph	Method	l	k				Method	l	k			
			3	5	7	10			3	5	7	10
SF50	Proposed method	1	0.4000	1.8400	3.4800	6.2400	Feder <i>et al.</i>	1	0.4000	1.7600	3.4400	6.2400
		2	6.5200	10.9200	14.4800	18.2000		2	6.6800	11.0000	14.5200	18.1600
		3	8.0400	14.1200	17.8000	22.1600		3	6.6800	13.1600	17.5200	22.1600
SF100	Proposed method	1	0.5600	2.1600	3.8400	6.5200	Feder <i>et al.</i>	1	0.5000	2.1000	3.8000	6.5000
		2	10.9000	16.7000	21.2600	27.0200		2	10.0600	16.3800	20.9400	27.0800
		3	14.2800	22.5600	27.7800	34.2800		3	12.5600	21.8600	12.4600	18.0800
SF200	Proposed method	1	0.5300	2.0600	3.8300	6.6400	Feder <i>et al.</i>	1	0.4500	1.9100	3.6600	6.4400
		2	12.2000	19.2000	25.6500	32.5400		2	11.1600	19.2700	25.9100	34.2600
		3	4.2000	8.3100	12.3800	18.2000		3	4.2000	8.3100	12.3800	18.2000
RA82	Proposed method	1	1.0244	2.9268	4.8537	7.8537	Feder <i>et al.</i>	1	0.9512	2.7317	4.7073	7.7073
		2	3.8780	7.7317	10.9268	15.0000		2	4.2683	7.3659	10.9512	14.8537
		3	3.9268	7.3171	11.1463	16.1707		3	3.5122	7.3415	11.3171	15.8780
RA129	Proposed method	1	0.8217	2.4031	4.3721	7.3643	Feder <i>et al.</i>	1	0.7287	2.3256	4.2636	7.2403
		2	5.5814	9.7674	14.2326	19.9070		2	5.4109	9.7674	14.4031	19.8140
		3	5.5814	10.0310	14.0620	21.3333		3	5.1783	9.9380	13.9690	21.1938
RA176	Proposed method	1	0.8523	2.4318	4.3636	7.3636	Feder <i>et al.</i>	1	0.7841	2.3977	4.3068	7.2955
		2	5.7386	9.7386	13.7614	19.5341		2	5.8977	9.8750	13.9091	19.0568
		3	5.7159	10.2273	13.7727	20.5000		3	5.6591	9.7386	14.2386	19.9091

Table 4

The amount of AVD error for different values of k and l and for the proposed method and for the Feder *et al.* [19] model applied to real-world graphs. In each experiment, the best value is highlighted

Graph	Method	l	k				Method	l	k			
			3	5	7	10			3	5	7	10
karate	Proposed method	1	0.4706	1.7059	3.2941	5.8824	Feder <i>et al.</i>	1	0.4118	1.6471	3.2941	5.8824
		2	5.1765	9.3529	11.7647	14.5294		2	4.4118	8.8235	11.5882	14.5294
		3	6.5882	11.0588	14.0000	17.5882		3	5.2353	11.4118	13.1176	17.6471
dwt_72	Proposed method	1	1.0000	2.9722	4.9722	7.9722	Feder <i>et al.</i>	1	0.9722	2.9167	4.9167	7.9167
		2	2.7222	6.7222	9.3889	12.8056		2	2.7222	6.7222	9.3611	12.7222
		3	2.7500	6.7222	9.5278	13.1111		3	2.7500	6.7500	9.47221	12.8889
lesmis	Proposed method	1	0.6234	1.5584	2.7532	4.8312	Feder <i>et al.</i>	1	0.5714	1.4805	2.4675	4.5195
		2	6.1299	11.2987	14.9351	19.7922		2	5.6883	10.6234	14.9351	19.9740
		3	4.2857	7.8961	18.1039	24.6753		3	11.7922	15.3766	19.6883	24.9091
can_96	Proposed method	1	0.0000	0.0000	1.3333	3.3333	Feder <i>et al.</i>	1	0.0000	0.0000	0.6667	3.0000
		2	3.5417	7.3750	10.4375	14.9792		2	3.4792	7.8333	10.7708	15.0625
		3	7.0000	10.6250	14.5417	19.1458		3	7.9375	12.1667	15.0000	18.5417
polbooks	Proposed method	1	0.0190	0.4000	1.3333	3.3333	Feder <i>et al.</i>	1	0.0190	0.2857	1.2000	3.2381
		2	6.7238	11.6381	15.5429	21.5238		2	5.1048	10.0952	15.4476	21.2571
		3	3.8667	7.9619	20.4190	27.2762		3	3.8667	7.9619	11.8286	28.2667
football	Proposed method	1	0.0000	0.0000	0.0000	0.1913	Feder <i>et al.</i>	1	0.0000	0.0000	0.0000	0.1217
		2	8.2783	14.0522	18.7652	25.4261		2	7.8783	13.8783	18.9565	25.6696
		3	4.7826	8.8000	12.5043	17.7043		3	4.7826	8.8000	12.5043	17.7043

Table 5

The amount of APL error for different values of k and l for the proposed method and for the Feder *et al.* [19] model applied to synthetic graphs. In each experiment, the best value is highlighted

Graph	Method	l	k				Method	l	k			
			3	5	7	10			3	5	7	10
SF50	Proposed method	1	0.0275	0.1472	0.3937	0.7244	Feder <i>et al.</i>	1	0.0672	0.4035	0.6476	0.9088
		2	0.8591	1.0615	1.1456	1.2346		2	0.8999	1.0672	1.1595	1.2354
		3	0.8950	1.1382	1.2223	1.3170		3	0.9562	1.1235	1.2191	1.3170
SF100	Proposed method	1	0.0185	0.0609	0.2154	0.4588	Feder <i>et al.</i>	1	0.0324	0.2507	0.4322	0.6449
		2	0.8718	1.0522	1.1225	1.1871		2	0.8441	1.0550	1.1229	1.1889
		3	0.9572	1.1209	1.1910	1.2602		3	0.9742	1.1188	1.0413	1.0980
SF200	Proposed method	1	0.0042	0.1035	0.3025	0.5887	Feder <i>et al.</i>	1	0.0674	0.3588	0.6084	0.8624
		2	1.1604	1.4144	1.5202	1.5894		2	1.1516	1.4265	1.5410	1.6154
		3	1.4687	1.4894	1.5098	1.5391		3	1.4687	1.4894	1.5098	1.5391
RA82	Proposed method	1	0.2143	1.4763	2.2697	2.9009	Feder <i>et al.</i>	1	1.6521	2.7352	3.1845	3.4299
		2	2.5621	3.1125	3.3643	3.6067		2	2.6834	3.1574	3.4176	3.6518
		3	2.4564	3.1098	3.4109	3.6729		3	2.4871	3.1833	3.5082	3.7199
RA129	Proposed method	1	0.1590	0.9227	1.5043	1.9891	Feder <i>et al.</i>	1	0.6486	1.5853	2.0178	2.3767
		2	2.0746	2.4457	2.6544	2.8300		2	2.0926	2.4628	2.6765	2.8623
		3	2.0110	2.4569	2.6622	2.8580		3	1.9873	2.4640	2.6766	2.8830
RA176	Proposed method	1	0.2011	0.9942	1.6827	2.2745	Feder <i>et al.</i>	1	0.8921	1.9239	2.4338	2.8156
		2	2.4295	2.8372	3.0393	3.2186		2	2.5297	2.8714	3.0653	3.2452
		3	2.3297	2.8523	3.0745	3.2801		3	2.4284	2.8485	3.0863	3.2912

- polbooks – a network of books about US politics, sold by Amazon. Edges in the network show the frequent purchases made by buyers. Data compiled by V. Krebs (www.orgnet.com);

- football – a network of American football games between Division IA colleges during the regular Fall 2000 season, as compiled by M. Girvan and M. Newman.

Table 6

The amount of APL error for different values of k and l for the proposed method and for the Feder *et al.* [19] model applied to real-world graphs. In each experiment, the best value is highlighted

Graph	Method	l	k				Method	l	k			
			3	5	7	10			3	5	7	10
karate	Proposed method	1	0.0566	0.0084	0.0682	0.3178	Feder <i>et al.</i>	1	0.0522	0.2892	0.4711	0.6333
		2	0.4871	0.7384	0.8293	0.9096		2	0.5816	0.7295	0.8276	0.9167
		3	0.5210	0.7937	0.9007	1.0094		3	0.6279	0.8097	0.8739	1.0112
dwt_72	Proposed method	1	1.7117	3.8509	4.6897	5.2770	Feder <i>et al.</i>	1	3.2516	4.2930	5.5074	5.9069
		2	4.3224	5.1851	5.5912	5.8576		2	4.3224	5.3146	5.5575	5.8580
		3	4.2660	5.2406	5.6358	5.9002		3	4.4167	5.3955	5.6260	5.9061
lesmis	Proposed method	1	0.0223	0.0112	0.0669	0.1923	Feder <i>et al.</i>	1	0.0361	0.1229	0.2418	0.3946
		2	0.4954	0.7254	0.8293	0.9390		2	0.6058	0.7189	0.8416	0.9387
		3	0.7500	0.7976	0.8724	0.9965		3	0.7965	0.8792	0.9127	1.0122
can_96	Proposed method	1	0.0459	0.0459	0.0935	0.1984	Feder <i>et al.</i>	1	0.0459	0.0459	1.0453	1.8569
		2	0.5907	1.7659	2.0946	2.3190		2	1.5547	2.0905	2.2089	2.3313
		3	1.3738	1.9106	2.2207	2.3995		3	1.9970	2.1569	2.2788	2.3786
polbooks	Proposed method	1	0.0264	0.0094	0.0304	0.1262	Feder <i>et al.</i>	1	0.0077	0.1170	0.4212	0.7243
		2	0.6234	0.9291	1.1033	1.2580		2	0.7908	1.0057	1.1390	1.2681
		3	1.1674	1.2068	1.1273	1.3097		3	1.1674	1.2068	1.2439	1.3441
football	Proposed method	1	0.0218	0.0218	0.0218	0.0151	Feder <i>et al.</i>	1	0.0218	0.0218	0.0218	0.0052
		2	0.4952	0.6642	0.7349	0.8026		2	0.5431	0.6897	0.7443	0.8050
		3	0.6218	0.6571	0.6896	0.7352		3	0.6218	0.6571	0.6896	0.7352

Table 7

The amount of ACC error for different values of k and l for the proposed method and for the Feder *et al.* [19] model applied to synthetic graphs. In each experiment, the best value is highlighted

Graph	Method	l	k				Method	l	k			
			3	5	7	10			3	5	7	10
SF50	Proposed method	1	0.0001	0.0005	0.0013	0.0026	Feder <i>et al.</i>	1	0.0003	0.0013	0.0022	0.0036
		2	0.0033	0.0046	0.0052	0.0059		2	0.0036	0.0046	0.0053	0.0059
		3	0.0036	0.0051	0.0058	0.0067		3	0.0040	0.0050	0.0058	0.0067
SF100	Proposed method	1	0.0000	0.0001	0.0003	0.0007	Feder <i>et al.</i>	1	0.0001	0.0004	0.0006	0.0010
		2	0.0015	0.0020	0.0022	0.0025		2	0.0015	0.0020	0.0022	0.0025
		3	0.0018	0.0023	0.0025	0.0028		3	0.0018	0.0022	0.0021	0.0023
SF200	Proposed method	1	0.0000	0.0000	0.0001	0.0003	Feder <i>et al.</i>	1	0.0000	0.0002	0.0003	0.0005
		2	0.0007	0.0010	0.0012	0.0013		2	0.0007	0.0010	0.0012	0.0013
		3	0.0011	0.0011	0.0012	0.0012		3	0.0011	0.0011	0.0012	0.0012
RA82	Proposed method	1	0.0001	0.0008	0.0015	0.0023	Feder <i>et al.</i>	1	0.0009	0.0020	0.0027	0.0033
		2	0.0018	0.0027	0.0032	0.0038		2	0.0020	0.0028	0.0033	0.0040
		3	0.0017	0.0027	0.0033	0.0040		3	0.0017	0.0028	0.0036	0.0042
RA129	Proposed method	1	0.0001	0.0004	0.0007	0.0011	Feder <i>et al.</i>	1	0.0002	0.0008	0.0012	0.0016
		2	0.0012	0.0017	0.0020	0.0024		2	0.0013	0.0017	0.0021	0.0024
		3	0.0012	0.0017	0.0020	0.0024		3	0.0012	0.0017	0.0021	0.0025
RA176	Proposed method	1	0.0000	0.0002	0.0005	0.0008	Feder <i>et al.</i>	1	0.0002	0.0006	0.0009	0.0011
		2	0.0009	0.0012	0.0014	0.0016		2	0.0009	0.0012	0.0014	0.0016
		3	0.0008	0.0012	0.0014	0.0016		3	0.0009	0.0012	0.0014	0.0017

5.2. Comparing the Proposed Method with the Feder Model

The following sections show the degree of utility distortion resulting from the modification of graph required to pro-

duce a (k, l) -anonymous graph. Tables 3, 5, and 7 show the amount of change in AVD, APL, and ACC of anonymous graphs in comparison with original ones for synthetic graph datasets, respectively. Table 3 shows the AVD error for anonymous graphs and compares it with correspond-

Table 8

The amount of ACC error for different values of k and l for the proposed method and for the Feder *et al.* [19] model applied to real-world graphs. In each experiment, the best value is highlighted

Graph	Method	l	k				Method	l	k			
			3	5	7	10			3	5	7	10
karate	Proposed method	1	0.0001	0.0003	0.0007	0.0022	Feder <i>et al.</i>	1	0.0006	0.0020	0.0034	0.0049
		2	0.0037	0.0062	0.0073	0.0084		2	0.0047	0.0062	0.0073	0.0086
		3	0.0041	0.0069	0.0085	0.0101		3	0.0052	0.0073	0.0082	0.0102
dwt_72	Proposed method	1	0.0005	0.0016	0.0024	0.0032	Feder <i>et al.</i>	1	0.0012	0.0020	0.0037	0.0046
		2	0.0020	0.0031	0.0039	0.0046		2	0.0020	0.0033	0.0038	0.0046
		3	0.0020	0.0032	0.0040	0.0047		3	0.0021	0.0035	0.0040	0.0047
lesmis	Proposed method	1	0.0000	0.0001	0.0002	0.0004	Feder <i>et al.</i>	1	0.0001	0.0003	0.0005	0.0009
		2	0.0012	0.0019	0.0023	0.0028		2	0.0015	0.0019	0.0024	0.0028
		3	0.0021	0.0023	0.0026	0.0032		3	0.0022	0.0026	0.0028	0.0032
can_96	Proposed method	1	0.0000	0.0000	0.0001	0.0001	Feder <i>et al.</i>	1	0.0000	0.0000	0.0008	0.0018
		2	0.0004	0.0017	0.0022	0.0028		2	0.0014	0.0022	0.0025	0.0028
		3	0.0011	0.0019	0.0025	0.0030		3	0.0021	0.0024	0.0027	0.0029
polbooks	Proposed method	1	0.0000	0.0000	0.0001	0.0002	Feder <i>et al.</i>	1	0.0000	0.0001	0.0005	0.0010
		2	0.0008	0.0014	0.0018	0.0022		2	0.0011	0.0016	0.0019	0.0023
		3	0.0020	0.0021	0.0019	0.0024		3	0.0020	0.0021	0.0023	0.0025
football	Proposed method	1	0.0000	0.0000	0.0000	0.0000	Feder <i>et al.</i>	1	0.0000	0.0000	0.0000	0.0000
		2	0.0009	0.0013	0.0015	0.0017		2	0.0010	0.0014	0.0015	0.0017
		3	0.0012	0.0014	0.0015	0.0017		3	0.0012	0.0014	0.0015	0.0017

Table 9

The average execution time (in seconds) for the proposed method and for the Feder *et al.* model

Graph	Execution time [s] for Feder <i>et al.</i> model				Execution time [s] for proposed method			
	Cost time	Solve time	Post-process time	Whole time	Cost time	Solve time	Post-process time	Whole time
karate	0.0007	25.0539	0.0722	30.5104	0.0362	4.9168	0.0760	11.9279
dwt_72	0.0000	3.5578	0.0152	7.3205	0.3596	3.6221	0.0162	7.5697
lesmis	0.0012	5137.9768	0.7115	5174.1992	0.3597	4491.9676	0.7061	4519.3890
can_96	0.0002	5047.8342	0.2352	5064.9843	0.5451	1833.5432	0.1329	1849.7912
polbooks	0.0001	5241.9255	0.8628	5292.6429	2.8282	5067.6278	1.8566	5146.3476
football	0.0036	6829.3870	1.6720	6840.1676	1.2615	6729.2509	1.9699	6740.8451
SF50	0.0003	995.9613	0.2244	1002.8781	0.1574	61.7497	0.1952	67.0347
SF100	0.0003	5613.8762	0.7388	5622.0445	2.3944	5130.4746	1.3076	5140.5596
SF200	0.0004	6936.9346	2.5641	6999.5473	8.6035	6820.4279	1.3090	6915.9520
RA82	0.0002	7.0836	0.0628	13.5097	0.4928	6.3735	0.03290	11.9961
RA129	0.0019	1727.7701	0.0915	1741.1626	4.3179	32.7018	0.2013	51.8183
RA176	0.0002	1795.2124	0.11818	1824.7315	8.9662	52.5578	0.1280	82.9223

ing original graphs for different k , l ($k \in \{3, 5, 7, 10\}$, $l \in \{1, 2, 3\}$) and for proposed method and for the Feder *et al.* model are shown. Similarly, Tables 5 and 7 are concerned with APL and ACC errors, while Tables 4, 6, and 8 are about AVD error, APL error, and ACC error of real-world graphs, respectively.

The results presented in Tables 3 and 4 show that, in most cases, the Feder *et al.* model achieves the best AVD, especially in all experiments for $l = 1$. This is rational, since the main objective of the model is to add a minimum number of edges to the original graph. Therefore, the (average)

degree of vertices are changes minimally. However, for the cases of $l > 1$, our implementation suggests an approximate approach that uses CV, which means that in some cases the proposed method may be the winner with respect to AVD error. Anyway, the relative differences are negligible. For example, AVD error of the proposed method for SF200, $k = 5$ and $l = 2$ equals to 19.2 while the value is 19.27 for the [19]. As expected, Tables 5–8 confirm the superiority of the proposed method in almost all cases in terms of APL and ACC errors. These values confirm that using a more precise cost function for adding new edges to

produce anonymous graphs, even for an approximated approach in the cases of $l > 1$, yields more useful datasets than the naïve approach used by [19]. These degree of superiority is more evident for APL errors, as the proposed method is optimized for this scenario. For instance, APL error for the proposed method is 0.1262 in polbooks for $k = 10$ and $l = 1$, while the value is 0.7243 for [19]. A similar discussion is true for ACC error, i.e. the proposed method achieves, usually, a result that is better than or equivalent to that shown in [19], since the proposed method attempts to introduce new edges that do not change the structural properties of the underlying original graph. In brief, the results confirm that the trivial idea of trying to only add a minimum number of edges does not necessarily achieve the best results, and that using a more elegant way to introduce new edges may produce more useful datasets.

5.3. Proposed Method Execution Time

In this section, the execution time of the proposed method is reported. The entire execution time consist of cost time, mathematical problem solution time, and post-process time. Table 9 shows these components of the proposed method's execution time. The results of the entire time columns confirm that the anonymization problem can be solved in a reasonable time frame, because it is usually considered as an offline problem. The entire time and solution time of the proposed method are better, in most cases, in comparison with the Feder *et al.* model. But, as far as the cost time columns are concerned, the Feder *et al.* model is quicker, in most cases, than the proposed method.

References

- [1] V. V. H. Pham, S. Yu, K. Sood, and L. Cui, "Privacy issues in social networks and analysis: a comprehensive survey", *IET Networks*, vol. 7, no. 2, pp. 74–84, 2018 (doi: 10.1049/iet-net.2017.0137).
- [2] B. Palanisamy, L. Liu, Y. Zhou, and Q. Wang, "Privacy-preserving publishing of multilevel utility-controlled graph datasets", *ACM Trans. on Internet Technol.*, vol. 18, no. 2, pp. 1–21, 2018 (doi: 10.1145/3125622).
- [3] P. Joshi and C. Kuo, "Security and privacy in online social networks: A survey", in *Proc. IEEE Int. Conf. on Multim. and Expo ICME 2011*, Barcelona, Spain, 2011 (doi: 10.1109/ICME.2011.6012166).
- [4] R. Gross and A. Acquisti, "Information revelation and privacy in online social networks", in *WPES '05: Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, Alexandria, VA, USA, 2005, pp. 71–80 (ISBN: 1-59593-228-3).
- [5] B. Krishnamurthy and C. Wills, "Characterizing privacy in online social networks", in *Proc. of the 1st Worksh. on Online Soc. Netw.*, Seattle, WA, USA, 2008, pp. 37–42 (doi: 10.1145/1397735.1397744).
- [6] T. Truta, M. Tsikerdekis, and S. Zeadally, "Privacy in social networks", in *Privacy in a Digital, Networked World*, S. Zeadally and M. Badra, Eds. Springer, 2015, pp. 263–289 (ISBN: 978-3-319-08470-1).
- [7] B. Zhou, J. Pei, and W. Luk, "A brief survey on anonymization techniques for privacy preserving publishing of social network data", *ACM SIGKDD Explor. Newslett.*, vol. 101 no. 2, pp. 12–22, 2008 (doi: 10.1145/1540276.1540279).
- [8] E. Zheleva and L. Getoor, "Preserving the privacy of sensitive relationships in graph data", in *Privacy, Security, and Trust in KDD. First ACM SIGKDD International Workshop, PinKDD 2007, San Jose, CA, USA, August 12, 2007, Revised Selected Papers*. Springer, 2008, pp. 153–171 (doi: 10.1007/978-3-540-78478-4_9).
- [9] R. Trujillo-Rasua and I. G. Yero, "k-metric antidimension: A privacy measure for social graphs", *Inform. Sciences*, vol. 328, pp. 403–417, 2016 (doi: 10.1016/j.ins.2015.08.048).
- [10] C.-H. Tai, P. S. Yu, D.-N. Yang, and M.-S. Chen, "Privacy-preserving social network publication against friendship attacks", in *Proc. 17th ACM SIGKDD Int. Conf. on Knowl. Discov. and Data Mining KDD'11*, San Diego, CA, USA, 2011, pp. 1262–1270 (doi: 10.1145/2020408.2020599).
- [11] W. Wentao, X. Yanghua, W. Wei, H. Zhenying, and W. Zhihui, "k-symmetry model for identity anonymization in social networks", in *Proc. 13th Int. Conf. on Ext. Database Technol. EDBT'10*, Lausanne, Switzerland, 2010, pp. 111–122 (doi: 10.1145/1739041.1739058).
- [12] J. Casas-Roma, J. Herrera-Joancomartí, and V. Torra, "An algorithm for k-degree anonymity on large networks", in *Proc. IEEE/ACM Int. Conf. on Adv. in Soc. Netw. Anal. and Mining ASONAM 2013*, Niagara Falls, ON, Canada, 2013, pp. 671–675 (doi: 10.1145/2492517.2492643).
- [13] T. Tassa and D. Cohen, "Anonymization of centralized and distributed social networks by sequential clustering", *IEEE Trans. on Knowl. and Data Engin.*, vol. 25, no. 2, pp. 311–324, 2013 (doi: 10.1109/TKDE.2011.232).
- [14] B. Fung, Y. Jin, J. Li, and J. Liu, "Anonymizing social network data for maximal frequent-sharing pattern mining", in *Recommendation and Search in Social Networks*, Ö. Ulusoy, A. Uz Tansel, and E. Arkun, Eds. Springer, 2015, pp. 77–100 (doi: 10.1007/978-3-319-14379-8_5).
- [15] H. Jiang, "A novel clustering-based anonymization approach for graph to achieve privacy preservation in social network", in *Proc. Int. Conf. on Adv. in Mechan. Engin. and Indust. Inform. AMEII 2015*, Zhengzhou, China, 2015, pp. 545–549 (doi: 10.2991/ameii-15.2015.102).
- [16] Z. Shiwen, L. Qin, and L. Yaping, "Anonymizing popularity in online social networks with full utility", *Future Gener. Comp. Syst.*, vol. 72, pp. 227–238, 2017 (doi: 10.1016/j.future.2016.05.007).
- [17] W. Yazhe, X. Long, B. Zheng, and K. C. Lee, "Utility-oriented k-anonymization on social networks", in *Database Systems for advanced Applications. 16th International Conference, DASFAA 2011, Hong Kong, China, April 22–25, 2011, Proceedings, Part I*, J. X. Yu, M. H. Kim, and R. Unland, Eds. Springer, 2011, pp. 78–92 (doi: 10.1007/978-3-642-20149-3_8).
- [18] C. Watanabe, T. Amagasa, and L. Liu, "Privacy risks and countermeasures in publishing and mining social network data", in *Proc. 7th Int. Conf. on Collab. Comput.: Network., Appl. and Worksharing CollaborateCom 2011*, Orlando, FL, USA, 2011 (doi: 10.4108/icst.collaboratecom.2011.247177).
- [19] T. Feder, S. U. Nabar, and E. Terzi, "Anonymizing graphs", arXiv:0810.5578 [cs.DB].
- [20] K. Stokes and V. Torra, "Reidentification and k-anonymity: a model for disclosure risk in graphs", *Soft Comput.*, vol. 16, no. 10, pp. 1657–1670, 2012 (doi: 10.1007/s00500-012-0850-4).
- [21] J. Cheng, A. W.-C. Fu, and J. Liu, "K-isomorphism: privacy preserving network publication against structural attacks", in *Proc. of the ACM SIGMOD Int. Conf. on Manag. of Data SIGMOD'10*, Indianapolis, Indiana, USA, 2010, 459–470 (doi: 10.1145/1807167.1807218).
- [22] K. Liu and E. Terzi, "Towards identity anonymization on graphs", in *Proc. ACM SIGMOD Int. Conf. on Manag. of Data SIGMOD'08*, Vancouver, Canada, 2008, pp. 93–106 (doi: 10.1145/1376616.3776629).
- [23] B. Zhou and J. Pei, "The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks", *Knowl. and Inform. Syst.*, vol. 28, no. 1, p. 47–77, 2011 (doi: 10.1007/s10115-010-0311-2).

- [24] L. Zou, L. Chen, and M. Özsu, "K-automorphism: a general framework for privacy preserving network publication", in *Proc. of the VLDB Endowment*, vol. 2, no. 1, pp. 946–957, 2009 (doi: 10.14778/1687627.1687734).
- [25] M. I. H. Ningga and J. H. Abawajy, "Utility-aware social network graph anonymization", *J. of Netw. and Comp. Appl.*, vol. 56, pp. 137–148, 2015 (doi: 10.1016/j.jnca.2015.05.013).
- [26] J. Casas-Roma, J. Herrera-Joancomartí, and V. Torra, "A survey of graph-modification techniques for privacy-preserving on network", *Artif. Intell. Rev.*, vol. 47, no. 3, pp. 341–366, 2017 (doi: 10.1007/s10462-016-9484-8).
- [27] M. Hay, G. Miklau, D. Jensen, P. Weis, and S. Srivastava, "Anonymizing social networks", Tech. Rep. no. 07-19, Computer Science Department, University of Massachusetts Amherst, 2007 [Online]. Available: https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1175;context=cs_faculty_pubs
- [28] K. Stokes and V. Torra, "On some clustering approaches for graphs", in *Proc. IEEE Int. Conf. on Fuzzy Syst. FUZZ-IEEE 2011*, Taipei, Taiwan, 2011, pp. 409–415 (doi: 10.1109/FUZZY.2011.6007447).
- [29] J. Casas-Roma, "Privacy-preserving on graphs using randomization and edge-relevance", in *Modeling Decisions for Artificial Intelligence 11th International Conference, MDAI 2014, Tokyo, Japan, October 29-31, 2014. Proceedings*, V. Torra, Y. Narukawa, Y. Endo, Eds. LNCS, vol. 8825. Springer, 2014, pp. 204–216 (doi: 10.1007/978-3-319-12054-6_18).
- [30] P. Samarati, "Protecting respondents' identities in microdata release", *IEEE Trans. Knowl. Data Engin. (TKDE)*, vol. 13, no. 6, pp. 1010–1027, 2001 (doi: 10.1109/69.971193).
- [31] L. Sweeney, "k-anonymity: a model for protecting privacy", *Int. J. of Uncert., Fuzziness Knowl.-Based Syst. (IJUFKS)*, vol. 10, no. 5, pp. 557–570, 2002 (doi: 10.1142/S0218488502001648).
- [32] B. Zhou and J. Pei, "Preserving privacy in social networks against neighborhood attacks", in *Proc. IEEE 24th Int. Conf. on Data Engin.*, Cancun, Mexico, 2008, pp. 506–515 (doi: 10.1109/ICDE.2008.4497459).
- [33] Y. Wang, L. Xie, B. Zheng, and K. C. K. Lee, "High utility k-anonymization for social network publishing", *Knowl. and Inform. Syst.*, vol. 41, no. 3, pp. 697–725, 2014 (doi: 10.1007/s10115-013-0674-2).
- [34] X. He, J. Vaidya, B. Shafiq, N. Adam, and V. Atluri, "Preserving privacy in social networks: a structure-aware approach", in *Proc. IEEE/WIC/ACM Int. Joint Conf. on Web Intell. and Intell. Agent Technol. WI-IAT'09*, Milan, Italy, 2009, pp. 47–54 (doi: 10.1109/WI-IAT.2009.108).
- [35] G. Kossinets and D. J. Watts, "Empirical analysis of an evolving social network", *Science*, vol. 311, no. 5757, pp. 88–90, 2006 (doi: 10.1126/science.1116869).
- [36] R. Mortazavi and S. H. Erfani, "An effective method for utility preserving social network graph anonymization based on mathematical modeling", *Int. J. of Engin.*, vol. 31, no. 10, pp. 1624–1632, 2018 [Online]. Available: http://www.ijeir.info/article_81694_e3849caf4384bbe53aed25f3afd8d93e.pdf
- [37] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks", *Proc. of the Nat. Acad. of Sci. USA*, vol. 99, no. 12, pp. 7821–7826, 2002 (doi: 10.1073/pnas.122653799).
- [38] L. J. Lu and M. Zhang, "Edge betweenness centrality", in *Encyclopedia of Systems Biology*, W. Dubitzky, O. Wolkenhauer, K.-H. Cho, H. Yokota, Eds. New York, NY: Springer, 2013, pp. 647–648 (doi: 10.1007/978-1-4419-9863-7_874).
- [39] CPLEX, GAMS. The solver manuals, GAME/CPLEX; 1996 [Online]. Available: https://www.gams.com/latest/docs/S_CPLEX.html



Seyedeh Hamideh Erfani received her B.Sc. degree and the M.Sc. degree from the Department of Computer Engineering, Ferdowsi University of Mashhad, Iran, in 2010 and 2013, respectively, from the Department of Computer Engineering, Ferdowsi University of Mashhad. She is a lecturer at Damghan University. Her main research

interests include computer vision, machine learning, and information systems.

E-mail: sh.erfani@du.ac.ir

School of Engineering

Damghan University

36716-41167, Damghan, Iran



Reza Mortazavi received his Ph.D. degree in Software Engineering in 2015 from Tarbiat Modares University, Iran. Currently he is an Assistant Professor at the School of Engineering, Damghan University, Iran. His current research interests include data privacy and machine learning.

E-mail: r.mortazavi@du.ac.ir

School of Engineering

Damghan University

36716-41167, Damghan, Iran

An Inclusive Performance Analysis of Single-branch Single-relay AF Transmission in a Mixed Rayleigh-Nakagami- m Fading Environment

Slimane Benmahmoud

Department of Electronics Engineering, Electrical Engineering Laboratory, M'sila University, M'sila, Algeria

<https://doi.org/10.26636/jtit.2019.129218>

Abstract—In this paper, the end-to-end performance of a single-branch two-hop amplify-and-forward (AF) relaying network in a mixed Rayleigh-Nakagami- m fading environment, is investigated. Four different fading scenarios and three standard relay configurations for each scenario are considered. Exact analytical expressions for the outage probability and tight upper bounds for the ergodic capacity are derived. Results of Monte Carlo simulations are provided to verify the accuracy of the analytical results.

Keywords—*amplify-and-forward (AF), decode-and-forward (DF), two/dual-hop relaying, outage probability, ergodic capacity, signal-to-noise ratio (SNR).*

1. Introduction

In a challenging wireless communication environment, cooperative relaying may be used to extend the coverage area and to enable high speed information transfers [1]–[4]. In relay networks, a source node (S) communicates with a destination node (D) through one or several intermediate nodes (R) referred to as relays. The two main relaying protocols are regenerative and non-regenerative relaying [5]. Decode-and-forward (DF) relaying [6], [7] is the most commonly used regenerative relaying protocol, where the relay detects and decodes the signal received from the source and then re-encodes and re-transmits it towards the destination. Amplify-and-forward (AF) relaying [8], [9] is a straightforward and popular non-regenerative relaying protocol. In AF relaying, the relay simply amplifies the signal received from the source and forwards it straight to the destination.

In a practical scenario, because of the geographical location of the relay node and its distances with respect to the source node and the destination node, different links may experience different fading conditions and, rarely, similar fading conditions. For example, one link may be in a line-of-sight (LOS) situation, while the other link may be in a non-LOS situation. Both of them may operate in a LOS or a non-LOS scenario as well. In the literature, the former situation

has been described as an asymmetric or mixed fading scenario, while the latter has been described as a symmetric fading scenario.

Initially, most works focusing on single-branch single-relay cooperative networks considered the symmetric scenario [8], [10]–[12]. Later, increasing interest in the asymmetric scenarios was observed. Authors of [3] derived exact and lower bound expressions for the outage probability and the average bit error probability for a two-hop AF relaying network over a mixed Rayleigh-Rician fading environment. After that and considering the same fading environment, more cooperative relaying models were studied, for instance the two-hop AF models in [13], [14] and the two-hop DF model in [15].

A mixed Nakagami-Rician fading environment was considered in [16], while the authors of [17] studied the performance of AF cooperative relaying in a mixed Rayleigh-generalized Gamma fading environment. The work [18] studied the performance two-hop AF relaying in an asymmetric Nakagami-lognormal fading environment, deriving novel closed-form expressions for the outage probability and the symbol error rate. In a recent study on two-hop AF relaying [19], a non-identical Rician fading environment was considered. Kumat *et al.* analyzed, in [20], the performance of QAM in AF cooperative networks over Rayleigh fading channels. In [21], the authors developed a unified framework to evaluate the performance of single-branch two-hop AF relaying, in the presence of transceiver hardware impairments, in a symmetric Nakagami- m fading environment.

A crucial design issue affecting AF relaying is the selection of the relay amplification factor based on the noise power and the source-to-relay channel state-information (CSI). Three standard AF relay configurations are known in the literature, namely channel-noise-assisted (CNA) relaying [22]–[24], channel-assisted (CA) relaying [11] and blind relaying [3], [12].

In the previously mentioned works and in numerous other studies related to the analysis of performance of single-branch dual-hop AF relaying over fading environments, ei-

ther their fading conditions are limited to one fading symmetric or asymmetric scenario, or their relay functioning is limited to just one configuration. In this study, four different fading scenarios are considered, where each of the two hops' links may experience either a Rayleigh or a Nakagami- m fading. Under each of the four fading scenarios, three different relay configurations are considered, namely CNA, CA and blind relaying. Such an approach renders twelve different fading scenario-relay configurations. The contribution of the paper consists in deriving an exact analytical expression for the outage probability and a tight upper bound for the ergodic capacity for each of the twelve cases referred to above.

The remainder of the paper is organized as follows. Section 2 presents the system and channel model. In Section 3, exact analytical expressions for the outage probability and tight upper bounds for the ergodic capacity are derived. Some Monte Carlo simulations are carried out and their results are provided in Section 4. Finally, Section 5 concludes the paper.

For quick reference, we wish to provide an explanation of the common notations used in this paper. $K_v(\cdot)$ is the modified Bessel function of the second kind [30] of order v . $W_{\mu,\nu}(\cdot)$ denotes the Whittaker function [30]. ${}_2F_1(\alpha, \beta; \gamma; \cdot)$ represents the Gauss hyper-geometric function [30]. $f_X(x)$, $F_X(x)$, $\bar{F}_X(x)$ and $M_X(s)$ denote PDF, CDF, complementary CDF and the MGF of a continuous RV X , respectively. Z^+ denotes the set of positive integers. The gamma function $\Gamma(n)$ of integer n satisfies $\Gamma(n) = (n-1)!$.

2. System and Channel Model

As shown in Fig. 1, this work considers a single-branch single-relay (i.e. a two-hop) AF cooperative relay network, where the transmission from source S to destination D via relay R takes place in two time slots. In the first time slot, S sends its signal to R . In the second time slot, R amplifies the signal received from S by a gain factor G , and forwards the resultant signal to D [19], [25], and [26]. It is assumed here that there is no direct or indirect path other than the $S \rightarrow R \rightarrow D$ path. For the end-to-end (E2E) signal-to-noise ratio (SNR) Λ of this system, the model proposed in [25] is considered:

$$\Lambda = \frac{\Lambda_1 \Lambda_2}{a \Lambda_1 + \Lambda_2 + b}, \quad (1)$$

where $\Lambda_i = \frac{P_i}{N_i} |h_i|^2$ represents the instantaneous SNR of the i -th hop, with P_i being the transmission power of s_i , N_i being the power of the additive white Gaussian noise (AWGN) component n_i .

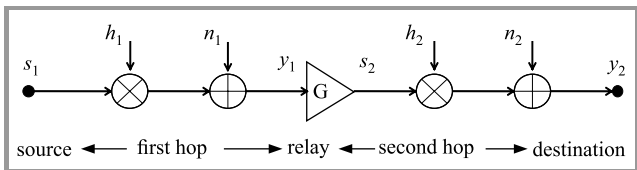


Fig. 1. Block diagram of single-branch dual-hop AF relaying.

The channel magnitude $\{|h_i|\}$ is modeled either as a Rayleigh or a Nakagami- m distributed RV, so as stated in Table 1, four different fading scenarios will be considered.

Table 1
Four different fading scenarios considered in this work

Case	$S-R$ link $\{ h_1 \}$	$R-D$ link $\{ h_2 \}$
Scenario 1	Nakagami- m	Nakagami- m
Scenario 2	Nakagami- m	Rayleigh
Scenario 3	Rayleigh	Nakagami- m
Scenario 4	Rayleigh	Rayleigh

The instantaneous SNR, Λ_i , is either exponential or Gamma distributed. Their PDFs are $f_{\Lambda_i}^{Exp}(x) = (1/\bar{\Lambda}_i) e^{-x/\bar{\Lambda}_i}; x \geq 0$, and $f_{\Lambda_i}^{\Gamma}(x) = (x^{\alpha_i-1}/\Gamma(\alpha_i)\beta_i^{\alpha_i}) e^{-x/\beta_i}; x \geq 0$, respectively.

The E2E SNR in Eq. (1), corresponds to a gain given by:

$$G^2 = \frac{P_2}{a|h_1|^2 P_1 + b N_1}, \quad (2)$$

where $a, b \geq 0$.

The setting $P_1 = P_2 = 1$ in Eq. (2) corresponds to the model considered in [25]. Note that the choice of the values of the parameters a and b reflects the relay configuration with $(a, b) \in \{(1, 1), (1, 0), (0, 1)\}$ representing CNA, CA, and the blind relay configuration, respectively.

3. Performance Analysis

Given the E2E SNR, Λ in Eq. (1), we provide here exact closed-form expressions for the outage probability and tight upper bounds for the ergodic capacity of the system in consideration. All the expressions will be given as functions of parameters a and/or b .

3.1. Outage Probability

The outage probability denoted by $P_{out}(x)$ is defined as the probability that the channel fading makes the effective end-to-end SNR fall below a certain threshold x characteristic of acceptable communication quality:

$$P_{out}(x) = \Pr(\Lambda \leq x) = \Pr\left[\frac{\Lambda_1 \Lambda_2}{a \Lambda_1 + \Lambda_2 + b} \leq x\right]. \quad (3)$$

After applying some algebraic manipulations, $P_{out}(x)$ may be expressed as:

$$P_{out}(x) = 1 - \int_0^\infty \bar{F}_{\Lambda_2}\left(ax + \frac{ax^2 + bx}{z}\right) f_{\Lambda_1}(z+x) dz \quad (4)$$

where $\bar{F}_{\Lambda_i}(x) = 1 - F_{\Lambda_i}(x)$ is the complementary CDF of Λ_i .

First Scenario:

- For the CNA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), will be given for $x \geq 0$ by

$$P_{out}(x) = 1 - 2e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{m=0}^k c_1(n, k, m) \times K_{n-m+1} 2\sqrt{\frac{(ax+b)x}{\beta_1\beta_2}} \left(a + \frac{b}{x}\right)^{\frac{n+m+1}{2}} \times x^{\alpha_1+k}, \quad (5)$$

$$\text{where } c_1(n, k, m) = \frac{a^{k-m} \beta_1^{\frac{n-m+1-2\alpha_1}{2}} \beta_2^{\frac{m-n-1-2k}{2}}}{m!(k-m)!n!(\alpha_1-1-n)!}.$$

- For the CA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), will be given for $x \geq 0$ by:

$$P_{out}(x) = 1 - 2e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{m=0}^k c_1(n, k, m) \times K_{n-m+1} 2\sqrt{\frac{ax}{\beta_1\beta_2}} a^{\frac{n+m+1}{2}} x^{\alpha_1+k}. \quad (6)$$

- For the blind relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ will be:

$$P_{out}(x) = 1 - 2e^{-\frac{1}{\beta_1}x} \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} c_2(n, k) \times K_{n-k+1} 2\sqrt{\frac{bx}{\beta_1\beta_2}} x^{\frac{2\alpha_1+k-n-1}{2}}, \quad (7)$$

$$\text{where } c_2(n, k) = \frac{\beta_1^{\frac{n-k+1-2\alpha_1}{2}} \left(\frac{b}{\beta_2}\right)^{\frac{n+k+1}{2}}}{n!k!(\alpha_1-n-1)!}.$$

Proof: See the appendix. ■

Second Scenario:

- For the CNA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), will be given for $x \geq 0$ by:

$$P_{out}(x) = 1 - 2e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} \sum_{n=0}^{\alpha_1-1} c_3(n) \times K_{n+1} 2\sqrt{\frac{(ax+b)x}{\beta_1\gamma_2}} \left(a + \frac{b}{x}\right)^{\frac{n+1}{2}} x^{\alpha_1}, \quad (8)$$

$$\text{where } c_3(n) = \frac{\beta_1^{\frac{n+1-2\alpha_1}{2}} \gamma_2^{-\frac{n+1}{2}}}{n!(\alpha_1-1-n)!}.$$

- For the CA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), will be given for $x \geq 0$ by:

$$P_{out}(x) = 1 - 2e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} \sum_{n=0}^{\alpha_1-1} c_3(n) \times K_{n+1} 2\sqrt{\frac{a}{\beta_1\gamma_2}} xa^{\frac{n+1}{2}} x^{\alpha_1}. \quad (9)$$

- For the blind relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ will be:

$$P_{out}(x) = 1 - 2e^{-\frac{1}{\beta_1}x} \sum_{n=0}^{\alpha_1-1} c_3(n) K_{n+1} 2\sqrt{\frac{bx}{\beta_1\gamma_2}} \times b^{\frac{n+1}{2}} x^{\frac{2\alpha_1-n-1}{2}}. \quad (10)$$

Proof: See the appendix. ■

Third Scenario:

- For the CNA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), will be given for $x \geq 0$ by:

$$P_{out}(x) = 1 - 2e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} \sum_{k=0}^{\alpha_2-1} \sum_{q=0}^k c_4(k, q) \times K_{1-q} 2\sqrt{\frac{(ax+b)x}{\gamma_1\beta_2}} \left(a + \frac{b}{x}\right)^{\frac{q+1}{2}} x^{k+1}, \quad (11)$$

$$\text{where } c_4(k, q) = \frac{a^{k-q} \gamma_1^{-\frac{q+1}{2}} \beta_2^{\frac{q-1-2k}{2}}}{q!(k-q)!}.$$

- For the CA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ will be:

$$P_{out}(x) = 1 - 2e^{-\left(\frac{1}{\gamma_1} + \frac{a}{\beta_2}\right)x} \sum_{k=0}^{\alpha_2-1} \sum_{q=0}^k c_4(k, q) \times K_{1-q} 2\sqrt{\frac{a}{\gamma_1\beta_2}} xa^{\frac{q+1}{2}} x^{k+1}. \quad (12)$$

- For the blind relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ will be:

$$P_{out}(x) = 1 - 2e^{-\frac{1}{\gamma_1}x} \sum_{k=0}^{\alpha_2-1} c_5(k) K_{1-q} 2\sqrt{\frac{bx}{\gamma_1\beta_2}} \times x^{\frac{k+1}{2}}, \quad (13)$$

$$\text{where } c_5(k) = \frac{\gamma_1^{-\frac{q+1}{2}} \beta_2^{\frac{b}{\beta_2} \frac{k+1}{2}}}{k!}.$$

Proof: See the appendix. ■

Fourth Scenario:

- For the CNA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ is:

$$P_{out}(x) = 1 - 2\sqrt{\frac{(ax+b)x}{\gamma_1\gamma_2}} e^{-\left(\frac{1}{\gamma_1} + \frac{a}{\gamma_2}\right)x} \times K_1 2\sqrt{\frac{(ax+b)x}{\gamma_1\gamma_2}}. \quad (14)$$

- For the CA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ will be:

$$P_{out}(x) = 1 - 2\sqrt{\frac{a}{\gamma_1\gamma_2}} x e^{-\left(\frac{1}{\gamma_1} + \frac{a}{\gamma_2}\right)x} \times K_1 2\sqrt{\frac{a}{\gamma_1\gamma_2}} x. \quad (15)$$

- For the CA relay configuration, the outage probability $P_{out}(x)$ in Eq. (4), for $x \geq 0$ will be:

$$P_{out}(x) = 1 - 2\sqrt{\frac{bx}{\gamma_1\gamma_2}} e^{-\frac{1}{\gamma_1}x} \times K_1 2\sqrt{\frac{a}{\gamma_1\gamma_2}} x. \quad (16)$$

Proof: See the appendix. ■

For $a = 1$, Eq. (15) reduces to Eq. (27) in [10].

For $b = c$, Eq. (16) reduces to Eq. (9) in [12].

3.2. The Ergodic Capacity

For a single-branch single-relay AF relaying network, the ergodic capacity (in bits/channel used) can be expressed as [26]:

$$C_{erg} \triangleq \frac{1}{2} E \{ \log_2(1 + \Lambda) \}, \quad (17)$$

where Λ is the end-to-end SNR in Eq. (1).

The ergodic capacity in Eq. (17) can be upper bounded as follows:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 + \int_0^\infty (P_{out}(x) - 1) dx \right). \quad (18)$$

Proof: See the appendix. ■

First Scenario:

- For the CNA relay configuration, the ergodic capacity C_{erg} in Eq. (18), will be:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{m=0}^k c_1(n, k, m) a^{\frac{n+m+1}{2}} \times \frac{\binom{n+m+2}{q} \left(\frac{b}{a}\right)^q}{\frac{\sqrt{a\beta_1\beta_2}\Gamma(n+2)\Gamma(m+1)}{2b(-1)^{\alpha_1+k-q+1}}} \times \frac{d^{\alpha_1+k-q+1}}{dp^{\alpha_1+k-q+1}} \left\{ e^{\frac{bp}{2a}} \times W_{-\frac{n+m+2}{2}, \frac{n-m+1}{2}} \frac{b(p - \sqrt{p^2 - \frac{4a}{\beta_1\beta_2}})}{2a} \times W_{-\frac{n+m+2}{2}, \frac{n-m+1}{2}} \frac{b(p + \sqrt{p^2 - \frac{4a}{\beta_1\beta_2}})}{2a} \right\}_{|p=\frac{1}{\beta_1} + \frac{a}{\beta_2}} \right). \quad (19)$$

- For the CA relay configuration, the ergodic capacity C_{erg} in Eq. (18), will be given by:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{m=0}^k c_1(n, k, m) a^{\frac{n+m+1}{2}} \times \frac{\sqrt{\pi}\Gamma(\alpha_1+k+n-m+2)\Gamma(\alpha_1+k-n+m)}{\Gamma(\alpha_1+k+3/2)} \times \left(\frac{16a}{\beta_1\beta_2}\right)^{\frac{n-m+1}{2}} \times \frac{{}_2F_1(\alpha_1+k+n-m+2; n-m+3/2; \alpha_1+k+3/2; \bar{s})}{\left(\left(\sqrt{\frac{a}{\beta_2}} + \sqrt{\frac{1}{\beta_1}}\right)^2\right)^{\alpha_1+k+n-m+2}} \right), \quad (20)$$

$$\text{where } \bar{s} = \left(\left(\sqrt{\frac{a}{\beta_2}} - \sqrt{\frac{1}{\beta_1}} \right)^2 / \left(\sqrt{\frac{a}{\beta_2}} + \sqrt{\frac{1}{\beta_1}} \right)^2 \right).$$

- For the blind relay configuration, the ergodic capacity C_{erg} in Eq. (18), will be given by:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} c_2(n, k) \times \frac{\Gamma(\alpha_1+1)\Gamma(\alpha_1+k-n)}{2\sqrt{\frac{b}{\beta_1\beta_2}}} \left(\frac{1}{\beta_1}\right)^{\frac{n-k-2k}{2}} \times e^{\frac{b}{2\beta_2}} W_{-\frac{n+m+2}{2}, \frac{n-m+1}{2}} \left(\frac{b}{\beta_2}\right) \right). \quad (21)$$

Proof: See the appendix. ■

Second Scenario:

- For the CNA relay configuration, the ergodic capacity C_{erg} in Eq. (18) will be:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{q=0}^{n+2} c_1(0,0,0) a^{\frac{n+1}{2}} \binom{n+2}{q} \times \left(\frac{b}{a} \right)^q \frac{\sqrt{a\beta_1\gamma_2} \Gamma(n+2) \Gamma(1)}{2b(-1)^{\alpha_1+k-q+1}} \frac{d^{\alpha_1-q+1}}{dp^{\alpha_1-q+1}} \left\{ e^{\frac{bp}{2a}} \times W_{-\frac{n+2}{2}, \frac{n+1}{2}} \frac{b(p - \sqrt{p^2 - \frac{4a}{\beta_1\gamma_2}})}{2a} \times W_{-\frac{n+2}{2}, \frac{n+1}{2}} \frac{b(p + \sqrt{p^2 - \frac{4a}{\beta_1\gamma_2}})}{2a} \right\}_{|p=\frac{1}{\beta_1} + \frac{a}{\gamma_2}} \right). \quad (22)$$

- For the CA relay configuration, the ergodic capacity C_{erg} in Eq. (18) may be derived from:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} c_3(n) a^{\frac{n+1}{2}} \times \frac{\sqrt{\pi} \Gamma(\alpha_1 + n + 2) \Gamma(\alpha_1 - n)}{\Gamma(\alpha_1 + 3/2)} \left(\frac{16a}{\beta_1\gamma_2} \right)^{\frac{n+1}{2}} \times \frac{{}_2F_1(\alpha_1 + n + 2; n + 3/2; \alpha_1 + 3/2; \bar{s})}{\left(\left(\sqrt{\frac{a}{\gamma_2}} + \sqrt{\frac{1}{\beta_1}} \right)^2 \right)^{\alpha_1 + n + 2}} \right). \quad (23)$$

- For the blind relay configuration, the ergodic capacity C_{erg} in Eq. (18) take the form:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} c_3(n) b^{\frac{n+1}{2}} \times \frac{\Gamma(\alpha_1 + 1) \Gamma(\alpha_1 - n)}{2 \sqrt{\frac{b}{\beta_1\gamma_2}}} \left(\frac{1}{\beta_1} \right)^{\frac{n-2\alpha_1}{2}} \times e^{\frac{b}{2\gamma_2}} W_{\frac{n-2\alpha_1}{2}, \frac{n+1}{2}} \frac{b}{\gamma_2} \right). \quad (24)$$

Proof: See the appendix. ■

Third Scenario:

- For the CNA relay configuration, the ergodic capacity C_{erg} in Eq. (18) is:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{q=0}^{m+2} c_1(0,k,m) a^{\frac{m+1}{2}} \times \binom{m+2}{q} \left(\frac{b}{a} \right)^q \frac{\sqrt{a\gamma_1\beta_2} \Gamma(2) \Gamma(m+1)}{2b(-1)^{2+k-q}} \times \frac{d^{k-q+2}}{dp^{k-q+2}} \left\{ e^{\frac{bp}{2a}} \times W_{-\frac{m+2}{2}, -\frac{m+1}{2}} \frac{b(p - \sqrt{p^2 - \frac{4a}{\gamma_1\beta_2}})}{2a} \times W_{-\frac{m+2}{2}, -\frac{m+1}{2}} \frac{b(p + \sqrt{p^2 - \frac{4a}{\gamma_1\beta_2}})}{2a} \right\}_{|p=\frac{1}{\gamma_1} + \frac{a}{\beta_2}} \right). \quad (25)$$

- For the CA relay configuration, the ergodic capacity C_{erg} in Eq. (18), will be given by:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} c_4(k,m) a^{\frac{m+1}{2}} \times \frac{\sqrt{\pi} \Gamma(k - m + 3) \Gamma(k + m + 1)}{\Gamma(k + 5/2)} \left(\frac{16a}{\gamma_1\beta_2} \right)^{\frac{1-m}{2}} \times \frac{{}_2F_1(k - m + 3; -m + 3/2; k + 5/2; \bar{s})}{\left(\left(\sqrt{\frac{a}{\beta_2}} + \sqrt{\frac{1}{\gamma_1}} \right)^2 \right)^{k-m+3}} \right). \quad (26)$$

- For the blind relay configuration, the ergodic capacity C_{erg} in Eq. (18) may be derived by:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{k=0}^{\alpha_2-1} c_5(k) \frac{\Gamma(2) \Gamma(k+1)}{2 \sqrt{\frac{b}{\gamma_1\beta_2}}} \left(\frac{1}{\gamma_1} \right)^{\frac{-k-2}{2}} e^{\frac{b}{2\beta_2}} W_{\frac{-k-2}{2}, \frac{-k+1}{2}} \frac{b}{\beta_2} \right). \quad (27)$$

Proof: See the appendix. ■

Fourth Scenario:

- For the CNA relay configuration, the ergodic capacity C_{erg} in Eq. (18), will be given by:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sum_{q=0}^2 c_1(0,0,0) a^{\frac{1}{2}} \binom{2}{q} \left(\frac{b}{a} \right)^q \times \frac{\sqrt{a\gamma_1\gamma_2} \Gamma(2) \Gamma(1)}{2b(-1)^{2-q}} \times \frac{d^{2-q}}{dp^{2-q}} \left\{ e^{\frac{bp}{2a}} W_{-1, \frac{1}{2}} \frac{b(p - \sqrt{p^2 - \frac{4a}{\gamma_1\gamma_2}})}{2a} \times W_{-1, \frac{1}{2}} \frac{b(p + \sqrt{p^2 - \frac{4a}{\gamma_1\gamma_2}})}{2a} \right\} \right)_{|p=\frac{1}{\gamma_1} + \frac{a}{\gamma_2}}. \quad (28)$$

- For the CA relay configuration, the ergodic capacity C_{erg} in Eq. (18), will be:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - 2 \sqrt{\frac{a}{\gamma_1\gamma_2}} \frac{\sqrt{\pi} \Gamma(3) \Gamma(1)}{\Gamma(5/2)} \left(\frac{16a}{\gamma_1\gamma_2} \right)^{\frac{1}{2}} \times \frac{{}_2F_1(3; 3/2; 5/2; \bar{s})}{\left(\left(\sqrt{\frac{1}{\gamma_1}} + \sqrt{\frac{a}{\gamma_2}} \right)^2 \right)^3} \right) = \frac{1}{2} \log_2 \left(1 - \frac{64}{3} \frac{a}{\gamma_1\gamma_2} \frac{{}_2F_1(3; 3/2; 5/2; \bar{s})}{\left(\sqrt{\frac{1}{\gamma_1}} + \sqrt{\frac{a}{\gamma_2}} \right)^6} \right). \quad (29)$$

- For the blind relay configuration, the ergodic capacity C_{erg} in Eq. (18) is:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 - \gamma_1 e^{\frac{b}{2\gamma_2}} W_{-1, \frac{1}{2}} \left(\frac{b}{\gamma_2} \right) \right). \quad (30)$$

Proof: See the appendix. ■

4. Numerical Results

In this section, we carry out some Monte Carlo simulations to verify the theoretical considerations. The four different scenarios, as mentioned previously, are used. For a Nakagami- m faded link, $\alpha_1 = \alpha_2 = 2$, and $\beta_1 = \beta_2 = 1$, while for a Rayleigh faded link, $\bar{\gamma} = 1$ was taken. We consider also the three standard relay configurations: CNA, CA and blind relaying.

First, in Fig. 2, we show the outage probability, $P_{out}(\lambda_{th})$, as a function of the average SNR for two different thresholds $\lambda_{th} = 2^2 - 1 = 3$ (i.e. 2 bits/channel used) and $\lambda_{th} = 2^5 - 1 = 31$ (i.e. 5 bits/channel used). From this figure, we can clearly see that the analytical expressions (represented by lines) match the results of Monte Carlo simulations (represented by symbols). For all four scenarios, CNA and CA relaying is characterized by almost the same outage probability which is higher than that of blind relaying, except for the second scenario where all three relay configurations perform almost exactly in the same manner.

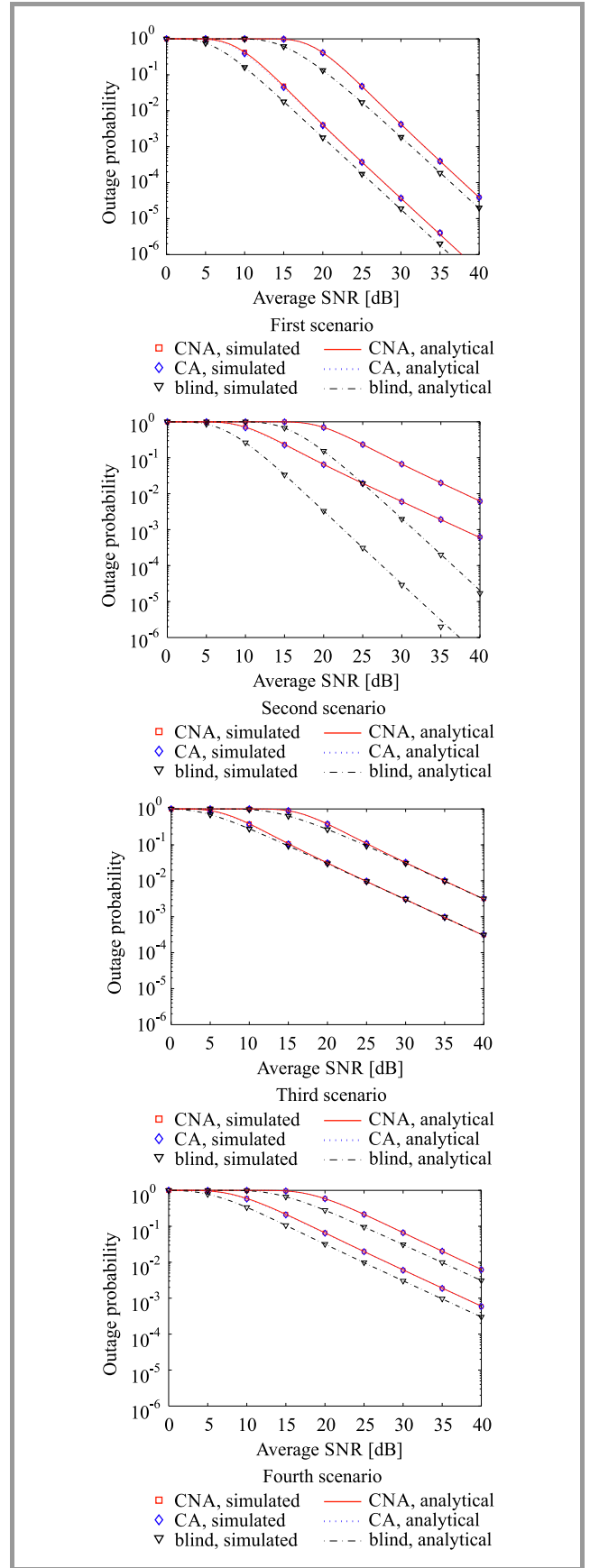


Fig. 2. Outage probability $P_{out}(\lambda_{th})$ for AF relaying; simulated and analytical results for CNA, CA, and blind relay configurations.

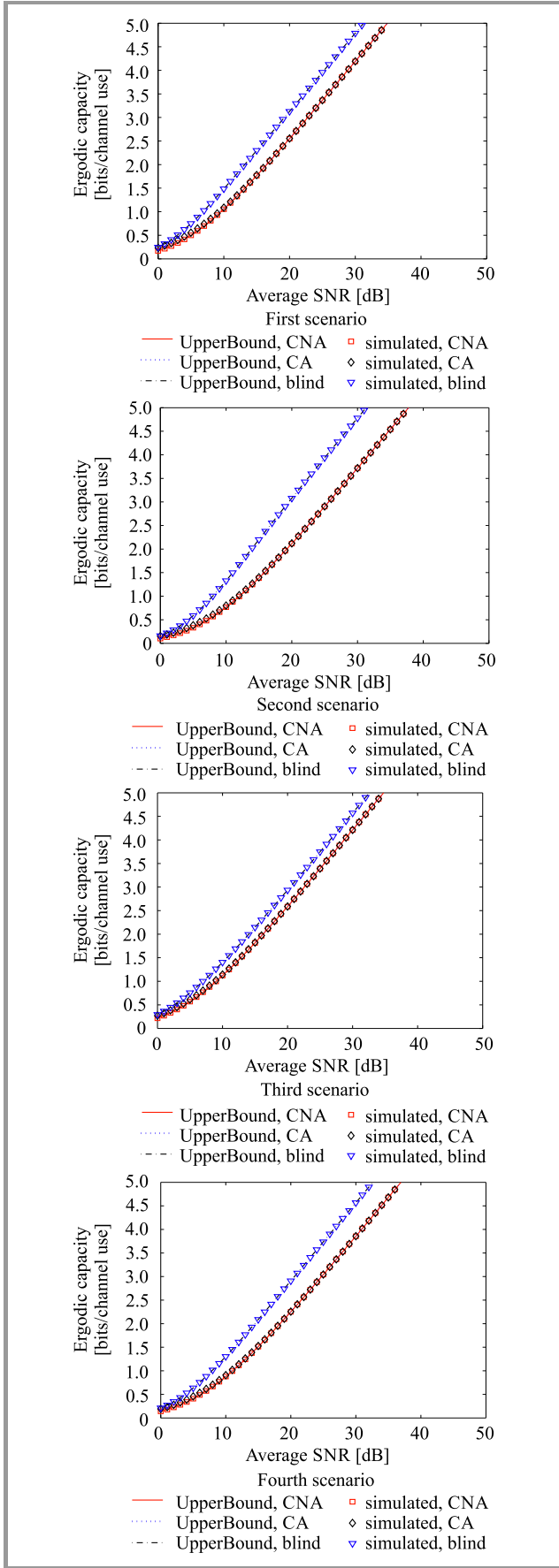


Fig. 3. Ergodic capacity C_{erg} for AF relaying; simulated and analytical results for CNA, CA, and blind relay configurations.

Figure 3 shows the ergodic capacity C_{erg} of AF relaying as a function of the average SNR. As shown, blind relaying, which suffered a lower outage probability, exhibits a higher ergodic capacity compared to CNA and CA relaying types. As a general comment, for both the outage probability and the ergodic capacity, the curves of CNA and CA relaying agree firmly. This approves the use of the former to approximate the latter in many applications.

5. Conclusion

In this paper, we derived novel closed-form expressions for the outage probability and tight upper bounds for the ergodic capacity of a single-branch two-hop relaying network under four different fading scenarios, with three standard relay configurations considered for each scenario. Monte Carlo simulations have been carried out to verify the accuracy of the analytical results.

Appendix

Proof of Results Shown in Section 3

Here, we provide the proof of results shown in Section 3. As stated in Section 2, the channel magnitude $\{|h_i|\}$ of the i -th hop is modeled either as a Rayleigh or a Nakagami- m distributed RV, so the instantaneous SNR, Λ_i , is to be modeled either as an exponential or a gamma distributed RV. Their CCDFs and PDFs are given for $i \in \{1, 2\}$ by:

$$\begin{cases} f_{\Lambda_i}^{Exp}(x) = \frac{1}{\lambda_i} e^{-\frac{x}{\lambda_i}}, & x \geq 0 \\ \bar{F}_{\Lambda_i}^{Exp}(x) = e^{-\frac{x}{\lambda_i}}, & x \geq 0. \end{cases} \quad (31)$$

$$\begin{cases} f_{\Lambda_i}^{Gamma}(x) = \frac{x^{\alpha_i-1} e^{-\frac{x}{\beta_i}}}{\Gamma(\alpha_i) \beta_i^{\alpha_i}}, & x \geq 0 \\ \bar{F}_{\Lambda_i}^{Gamma}(x) = \sum_{k=0}^{\alpha_i-1} \frac{e^{-\frac{x}{\beta_i}}}{k!} \left(\frac{x}{\beta_i}\right)^k, & x \geq 0 \end{cases} \quad (32)$$

1. Proof of the outage probability results (Subsection 3.1)

Let $Pr\{\Lambda\}$ denote the probability of event Λ . The outage probability $P_{out}(x)$ can be derived as:

$$\begin{aligned} P_{out}(x) &= \Pr \left\{ \frac{\lambda_1 \lambda_2}{a\lambda_1 + \lambda_2 + b} \leq x \right\} \\ &= \int_0^x \Pr \left\{ \lambda_2 \geq \frac{(a\lambda_1 + b)x}{\lambda_1 - x} \right\} \cdot f_{\lambda_1}(\lambda_1) d\lambda_1 \\ &\quad + \int_x^\infty \Pr \left\{ \lambda_2 \leq \frac{(a\lambda_1 + b)x}{\lambda_1 - x} \right\} \cdot f_{\lambda_1}(\lambda_1) d\lambda_1 \\ &= 1 - \int_x^\infty \bar{F}_{\lambda_2} \left(\frac{(a\lambda_1 + b)x}{\lambda_1 - x} \right) \cdot f_{\lambda_1}(\lambda_1) d\lambda_1. \end{aligned} \quad (33)$$

Simplifying with $z = \lambda_1 - x$, leads to Eq. (4).

First Scenario:

CNA case: substituting from Eq. (32) in Eq. (33) gives:

$$P_{out}(x) = 1 - \sum_{k=0}^{\alpha_2-1} \frac{e^{\left(\frac{ax}{\beta_2} + \frac{x}{\beta_1}\right)}}{k! \beta_2^k \Gamma(\alpha_1) \beta_1^{\alpha_1}} (ax^2 + bx)^k$$

$$= 1 - \int_0^\infty (z+x)^{\alpha_1-1} \left(\frac{1}{z} + \frac{ax}{ax^2 + bx} \right)^k$$

$$\times e^{-\left(\frac{(ax^2+bx)/\beta_2}{z} + \frac{z}{\beta_1}\right)} dz. \quad (34)$$

The use of lemma 4 from [26] yields Eq. (5) after some simplifications.

CA case: Eq. (6) can be directly deduced from Eq. (5) by setting $b = 0$.

Blind case: By setting $a = 0$ Eq. (34) reduces to:

$$P_{out}(x) = 1 - \frac{e^{-\frac{x}{\beta_1}} x^{\alpha_1-1}}{\Gamma(\alpha_1) \beta_1^{\alpha_1}} \sum_{k=0}^{\alpha_2-1} \frac{\left(\frac{x}{\beta_2}\right)^k}{k!}$$

$$\times \int_0^{+\infty} \left(\frac{b}{z}\right)^k \left(1 + \frac{z}{x}\right)^{\alpha_1-1} e^{-\left(\frac{z}{\beta_1} + \frac{bx}{\beta_2 z}\right)} dz. \quad (35)$$

The binomial expansion of the second term in the integral and the use of [28, 3.471.9] yields Eq. (7).

Second Scenario:

CNA case: substituting from Eq. (31) and Eq. (32) in Eq. (4) gives:

$$P_{out}(x) = 1 - \frac{e^{-\left(\frac{a}{\beta_2} + \frac{1}{\beta_1}\right)}}{\Gamma(\alpha_1) \beta_1^{\alpha_1}}$$

$$\times \int_0^\infty (z+x)^{\alpha_1-1} e^{-\left(\frac{(ax^2+bx)/\beta_2}{z} + \frac{z}{\beta_1}\right)} dz. \quad (36)$$

The binomial expansion of the term $(z+x)^{\alpha_1-1}$ and the use of [28, 3.471.9] yields Eq. (8) after some simplifications.

CA and blind cases: Eqs. (9) and (10) can be directly deduced from Eq. (8) by setting $b = 0$ and $a = 0$, respectively.

Third Scenario:

CNA case: substituting from Eqs. (31) and (32) in Eq. (4) gives:

$$P_{out}(x) = 1 - \sum_{k=0}^{\alpha_2-1} \frac{e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x}}{k! \beta_2^k \gamma_1} (ax^2 + bx)^k$$

$$\times \int_0^{+\infty} \left(\frac{ax}{ax^2 + bx} + \frac{1}{z}\right)^k e^{-\left(\frac{(ax^2+bx)/\beta_2}{z} + \frac{z}{\beta_1}\right)} dz. \quad (37)$$

The binomial expansion of the term $\left(\frac{ax}{ax^2+bx} + \frac{1}{z}\right)^k$ and the use of [28, 3.471.9] yields Eq. (11) after some simplifications.

CA case: Eq. (12) can be directly deduced from Eq. (11) by setting $b = 0$.

Blind case: by setting $a = 0$, Eq. (37) reduces to:

$$P_{out}(x) = 1 - \sum_{k=0}^{\alpha_2-1} \frac{e^{-\frac{x}{\gamma_1}}}{k! \beta_2^k \gamma_1} (bx)^k$$

$$\times \int_0^{+\infty} \left(\frac{1}{z}\right)^k e^{-\left(\frac{bx/\beta_2}{z} + \frac{z}{\gamma_1}\right)} dz. \quad (38)$$

The use of [28, 3.471.9] to evaluate the integral in Eq. (38), yields Eq. (13).

Fourth Scenario:

CNA case: substituting from Eq. (31) in Eq. (4) gives:

$$P_{out}(x) = 1 - \int_0^\infty e^{-\frac{1}{\beta_2} \left(ax + \frac{ax^2+bx}{z}\right)} \frac{1}{\gamma_1} e^{-\frac{1}{\gamma_1} (z+x)} dz. \quad (39)$$

The use of [28, 3.324.1] yields Eq. (14) after some simplifications.

CA and blind cases: Eqs. (15) and (16) can be directly deduced from Eq. (14) by setting $b = 0$ and $a = 0$, respectively.

2. Proof of ergodic capacity results (Subsection 3.2)

Let $E\{\Lambda\}$ denote the expected value of event Λ . The ergodic capacity C_{erg} can be expressed as in Eq. (17).

We note that the function $\log_2(1 + \Lambda)$ in Eq. (17), is twice-differentiable and its second derivative is:

$$-\frac{1}{\ln(2)(1 + \Lambda)^2} < 0. \quad (40)$$

Jensen's inequality could be then applied to get:

$$C_{erg} \leq \frac{1}{2} \log_2 \left(1 + E\{\Lambda\}\right). \quad (41)$$

We know that the expected value $E\{\Lambda\}$ of an event Λ is linked to its moment-generating function by:

$$E\{\Lambda\} = \frac{d}{ds} M_\Lambda(s)|_{s=0}. \quad (42)$$

Substituting Eq. (6) from [25] in Eq. (42) and substituting the result in Eq. (41) yields Eq. (18).

First Scenario:

CNA case: substituting from Eq. (5) in Eq. (18) gives:

$$C_{erg} \leq \frac{1}{2} \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{m=0}^k c_1(n, k, m)\right.$$

$$\times \int_0^{+\infty} e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} \left(a + \frac{b}{x}\right)^{\frac{n+m+1}{2}}$$

$$\times x^{\alpha_1+k} K_{n-m+1} 2 \sqrt{\frac{x(ax+b)}{\beta_1 \beta_2}} dx \Big). \quad (43)$$

From the integral in Eq. (43), we separate factor $(a + (b/x))^{n+m+2}$ and we perform a binomial expansion on it. Then we re-arrange the result into a Laplace transformation that can be evaluated and simplified using the identities shown in [29, 4.1.6] and [29, 4.17.20] to yield Eq. (19).

CA case: substituting from Eq. (6) in Eq. (18) gives:

$$C_{erg} \leq \frac{1}{2} \log \left(1 - 2 \sum_{n=0}^{\alpha_1-1} \sum_{k=0}^{\alpha_2-1} \sum_{m=0}^k c_1(n, k, m) \times \int_0^{+\infty} e^{-\left(\frac{1}{\beta_1} + \frac{a}{\beta_2}\right)x} a^{\frac{n+m+1}{2}} x^{\alpha_1+k} \times K_{n-m+1} 2\sqrt{\frac{a}{\beta_1\beta_2}} x dx \right). \quad (44)$$

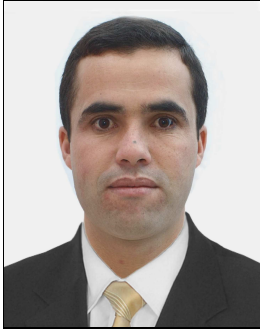
Evaluating the integral in Eq. (44) using [28, 6.621.3] yields Eq. (20).

Blind case: substituting from Eq. (7) in Eq. (18) and applying [28, 6.621.3] yields Eq. (21).

NB. The results in three other scenarios may be proven by following exactly the same steps as shown for the first scenario.

References

- [1] X. Xiao and L. J. Cimini, Jr., "Impact of overhead on spectral efficiency of cooperative relaying", *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2228–2239, 2013 (doi: 10.1109/TWC.2013.040413.120779).
- [2] A. Nostratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks", *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, 2004 (doi: 10.1109/MCOM.2004.1341264).
- [3] H. A. Suraweera, R. H. Y. Louie, Y. Li, G. K. Karagiannis, and B. Vucetic, "Two hop amplify-and-forward transmission in mixed Rayleigh and Rician fading channels", *IEEE Commun. Lett.*, vol. 13, no. 4, pp. 227–229, 2009 (doi: 10.1109/LCOMM.2009.081943).
- [4] Y. Kim and H. Liu, "Infrastructure relay transmission with cooperative MIMO", *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2180–2188, 2008 (doi: 10.1109/TVT.2007.912962).
- [5] M. Dohler and Y. Li, *Cooperative Communications: Hardware, Channel and PHY*. Wiley, 2010 (ISBN: 978-0-470-99768-0).
- [6] S. Ikki and M. H. Ahmed, "Performance of decode and forward cooperative diversity networks over Nakagami- m fading channels", in *Proc. IEEE Global Telecommun. Conf. IEEE GLOBECOM 2007*, Washington, DC, USA, 2007, pp. 4328–4333 (doi: 10.1109/GLOCOM.2007.823).
- [7] K. Chaitanya and P. Muthuchidambaramanathan, "Performance analysis of decode and forward relaying over dual-hop mixed fading channels", *Int. J. of Electron. Commun. (AEU)*, vol. 73, pp. 84–88, 2017 (doi: 10.1016/j.aeu.2016.12.026).
- [8] M. O. Hasna and M. S. Alouini, "End-to-end performance of transmission systems with relays over Rayleigh fading channels", *IEEE Trans. Wireless. Commun.*, vol. 2, no. 6, pp. 1126–1131, 2003 (doi: 10.1109/TWC.2003.819030).
- [9] N. Kumar, V. Bhatia, and D. Dixit, "Performance analysis of QAM in amplify-and-forward cooperative communication networks over Rayleigh fading channels", *Int. J. of Electron. Commun. (AEU)*, vol. 72, pp. 86–94, 2017 (doi: 10.1016/j.aeu.2016.11.022).
- [10] M. O. Hasna and M. S. Alouini, "End-to-end outage probability of multi-hop transmission over lognormal shadowed channels", *Arab. J. Sci. and Engin.*, vol. 28, no. 2C, pp. 35–44, 2003 [Online]. Available: <https://pdfs.semanticscholar.org/e7bc/4f4149251749702f6a61dc3e3fb53f56a485.pdf>
- [11] M. O. Hasna and M. S. Alouini, "Performance analysis of two-hop relayed transmissions over Rayleigh fading channels", in *Proc. IEEE 56th Veh. Technol. Conf.*, Vancouver, BC, Canada, 2002, vol. 4, pp. 1992–1996 (doi: 10.1109/VETECF.2002.1040567).
- [12] M. O. Hasna and M. S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays", *IEEE Trans. Wirel. Commun.*, vol. 3, no. 6, pp. 1963–1968, 2004 (doi: 10.1109/TWC.2004.837470).
- [13] S. H. Chen, F. Liu, X. Zhang, Y. N. Han, and D. C. Yang, "On the performance of two-hop amplify and forward relay networks with beamforming over Rayleigh-Rician fading channels", in *Proc. IEEE 72nd IEEE Veh. Technol. Conf. Fall (VTC 2010-Fall)*, Ottawa, Canada, 2010, pp. 1–4 (doi: 10.1109/VETECF.2010.5594283).
- [14] J. Ouyang, M. Lin, and Y. Zhuang, "Performance analysis of cooperative relay networks over asymmetric fading channels", *Electron. Lett.*, vol. 48, no. 21, pp. 1370–1371, 2012 (doi: 10.1049/el.2012.1737).
- [15] C. Q. Yang, W. B. Wang, S. A. Zhao, and M. G. Peng, "Opportunistic decode-and-forward cooperation in mixed Rayleigh and Rician fading channels", *ETRI J.*, vol. 2, pp. 287–290, 2011 (doi: 10.42187/etrij.11.0210.0219).
- [16] W. Xu, J. Zhang, and P. Zhang, "Performance analysis of dual-hop amplify-and-forward relay system in mixed Nakagami- m and Rician fading channels", *Electron. Lett.*, vol. 46, no. 17, pp. 1231–1232, 2010 (doi: 10.1049/el.2010.1555).
- [17] K. Nuri, B. Mehmet, and D. Ibrahim, "SER performance of amplify-and-forward cooperative diversity over asymmetric fading channels", *Wirel. Pers. Commun.*, vol. 73, no. 3, pp. 1117–1127, 2013 (doi: 10.1007/s11277-013-1251-4).
- [18] W. Cheng, "On the performance of dual-hop amplify-and-forward transmission in asymmetric multipath/shadowing fading channels", *Inform. Technol. J.*, vol. 12, pp. 4308–4312, 2013 (doi: 10.3923/itj.2013.4308.4312).
- [19] M. Delibasic and M. Pejanovic-Djurisic, "Dual-hop amplify-and-forward relay system over non-identical Rician fading channels", *Wirel. Pers. Commun.*, vol. 93, pp. 675–686, 2017 (doi: 10.1007/s11277-014-2221-1).
- [20] N. Kumar, V. Bhatia, and D. Dixit, "Performance analysis of QAM in amplify-and-forward cooperative communication networks over Rayleigh fading channels", *Int. J. Electron. Commun. (AEU)*, vol. 72, pp. 86–94, 2017 (doi: 10.1016/j.aeu.2016.11.022).
- [21] S. Benmhamoud and H. Dai, "A unified framework for performance evaluation of single-branch dual-hop AF relaying under the presence of transceiver impairments", *Physical Commun. J.*, vol. 32, pp. 231–241, 2019 (doi: 10.1016/j.phycom.2018.12.008).
- [22] J. N. Laneman and G. W. Wornell, "Energy-efficient antenna sharing and relaying for wireless networks", in *Proc. IEEE Wirel. Commun. and Network. IEEE WCNC*, Chicago, IL, USA, 2000, vol. 1, pp. 7–12 (doi: 10.1109/WCNC.2000.904590).
- [23] R. H. Y. Louie, Y. Li, and B. Vucetic, "Performance analysis of beamforming in two hop amplify and forward relay networks", in *Proc. IEEE Int. Conf. on Commun. ICC 2008*, Beijing, China, 2008, pp. 4311–4315 (doi: 10.1109/ICC.2008.809).
- [24] C. K. Datsikas, K. P. Peppas, F. I. Lazarakis, and G. S. Tombras, "Error rate performance analysis of dual-hop relaying transmissions over generalized- K fading channels", *Int. J. Electron. Commun. (AEU)*, vol. 64, no. 11, 2010, pp. 1094–1099 (doi: 10.1016/j.aeu.2009.09.005).
- [25] D. Senaratne and C. Tellambura, "Unified exact performance analysis of two-hop amplify-and-forward relaying in Nakagami fading", *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1529–1534, 2010 (doi: 10.1109/TVT.2009.2038784).
- [26] E. Bjornson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying: Performance limits with hardware impairments", *IEEE Trans. on Commun.*, vol. 61, no. 11, pp. 4512–4525, 2013 (doi: 10.1109/TCOMM.2013.100913.130282).
- [27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic Press, 2007 (ISBN 9780123736376).
- [28] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Tables of Integral Transforms*, vol. 1. New York: McGraw-Hill, 1954 (doi: 10.1002/zamm.19540341220).
- [29] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1970 (ISBN 978-0486612720).



sity (NCSU), USA, where he conducted research on

Slimane Benmahmoud received his B.Sc. from M'sila University, the M.Sc. and Ph.D. degrees from Djillali Liabes University in Algeria. Between 2010 and 2012, he was employed as an instructor at El-oued University in Algeria. From 2014 to 2016 he was a visiting research scholar at North Carolina State University

convex optimization and cooperative communications. He is currently a lecturer at the Department of Electronics Engineering at M'sila University. He is also a member of the Electrical Engineering Laboratory (LGE) at the same university. His main research interests are in coded modulation and iterative decoding, convex optimization and cooperative communications.

 <https://orcid.org/0000-0003-3778-5799>

E-mail: slimane34a@gmail.com

Department of Electronics Engineering

M'sila University

M'Sila 28000, Algeria

New OSTBC for Blind Channel Estimation and Tracking in MIMO-OFDM Systems

Abdelkader Tami¹, Mokhtar Keche², and Boubaker S. Bouazza¹

¹ *Laboratory Technology of Communication, Faculty of Technology, University of Saida, Algeria*

² *Laboratory Signal and Image, Department of Electronics, University of Sciences and Technology Oran, Algeria*

<https://doi.org/10.26636/jtit.2019.133819>

Abstract—Applying orthogonal space time block coding (OSTBC) to multiple-input multiple-output (MIMO) systems helps reduce receiver complexity. However, this approach has been applied only to flat fading channels, as when the channel is a frequency selective fading MIMO channel, OSTBC cannot be used directly since its orthogonal propriety may be lost. Furthermore, the MIMO channel is not always known. To deal with this problem, many techniques were proposed to estimate the channel using a training sequence. Unfortunately, these techniques reduce the useful spectral bandwidth. This paper proposes OSTBC for blind channel estimation and data detection in the case of a MIMO frequency selective channel. The aim of this new OSTBC is twofold: to solve the ambiguity of channel estimation and to reduce the complexity of the detector. By exploiting the well-known technique of orthogonal frequency division multiplexing (OFDM), the frequency selective fading MIMO channel is split into a set of flat fading sub-channels. Moreover, to accommodate the fact that a MIMO channel can be time varying, the steady state Kalman channel estimator (SS-KCE) is extended to track the channel's fast variations. The performance of the proposed blind algorithm is related by the adequate choice of the number of subcarriers and it is compared with other existing approaches by means of Monte Carlo simulations.

Keywords—MIMO, OFDM, blind channel estimation, OSTBC, SS-KCE.

1. Introduction

The growing interest in wireless mobile communication requires high performance and quality communication. This pushes researchers to develop many techniques with the aim of increasing data transmission rates. Multi-input multi-output (MIMO) is one of the techniques proposed to boost useful channel capacity without increasing total transmission power [1]. In addition, space-time block coding (STBC) has appeared as a powerful approach to exploit spatial diversity and to combat fading in MIMO wireless channels [2], [3].

The higher the data rate, the more the wireless channel becomes frequency selective and the more Inter Symbol

Interference (ISI) is encountered. To eliminate the ISI, orthogonal frequency division multiplexing (OFDM) is used in many standards, such as those pertaining to broadband wireless metropolitan area networks (IEEE 802.16) and mobile broadband wireless access networks (IEEE 802.20) [4]. The OFDM technique has some drawbacks too, such as its peak-to-average power ratio (PAPR) [5], but it still remains a candidate for 5G after introducing some modifications, such as filter bank multiple carrier (FBMC) [6] or generalized frequency division multiplexing (GFDM) [6]. The basic idea behind the OFDM technique consists in converting a frequency selective channel into a bank of flat fading sub-channels [7]. In addition, a combination between orthogonal STBC and OFDM is sought, because it allows to reach the maximum transmit diversity and simplifies the receiver structure [8].

Besides being time variable, a wireless channel may also be unknown to both the receiver and the transmitter. Generally, a periodic training sequence is transmitted to identify this wireless channel, but, unfortunately, this sequence occupies nearly 20% of the useful data rate. Here, in order to identify a wireless channel, we focus our interest on the blind channels which require no training pilot. Many research works dealing with channel estimation in MIMO-OFDM systems can be cited. For instance, paper [9] proposed a subspace blind channel estimation method for MIMO-OFDM systems, based on frequency correlation between adjacent subcarriers. Also, Sarmadi *et al.* in [10] proposed a semi-definite relaxation technique to achieve blind channel estimation for orthogonal STBC-OFDM systems. Another method was also proposed in [11]. It is based on the weighted covariance approach which explores all redundancies contained in the orthogonal code to resolve channel estimation non-scalar ambiguities. In [12], Zhang *et al.* developed a new subspace channel estimation method for OSTBC with MIMO-OFDM, using the block cyclic property of the channel matrix and the cyclic repetition method (CRM). However, this method needs a few pilots to alleviate channel estimation ambiguity. A new approach was developed by Jiang *et al.* in [8], where the

Table 1
List of abbreviations

A	Diagonal matrix controls channel coefficients
$\hat{B}^{(\ell,q)}$	Block contains codeword symbols corresponding to channel estimating matrix $\hat{\mathcal{H}}^{(\ell)}[u]$
$\{C\}$	Set of modulation constellation points
F	Transition diagonal matrix
f_0	Carrier frequency
$F_D T_s$	Normalized Doppler rate
G	Generator matrix
g	Discrete modulation waveform
H	Channel matrix of a frequency Rayleigh selective fading
$\mathcal{H}$	Orthogonal matrix groups channel attenuation effect and coding effect
$\hat{\mathcal{H}}$	Estimated matrix channel
$\mathbf{h}_{rt}$	Sub-channel between the t -th transmitting antenna and the r -th receiving antenna
$\hat{h}_{rj}$	FFT of the sub-channel $\mathbf{h}_{rj}$
$\hat{h}$	Vector of the matrix $\mathcal{H}$
$I_{N_r N_t}$	$N_r \times N_t$ unitary matrix
J_0	Zero-order Bessel function of the first kind
k	Block OFDM number
L	Frequency sub-channel length
ℓ, ℓ_1, ℓ_2 and n	Integer numbers
m	Total number of constellation points
N	OFDM subcarrier number
N_g	Length of cyclic prefix
N_t	Transmitting antennas
N_r	Receiving antennas
n_s	Energy of each uncoded data block
P_∞	Steady state estimation error covariance matrix
$\text{Re}(\cdot)$	Real part
s_i	Transmitted symbol with $i = 1, 2, \dots, 4N$
$\hat{s}$	Estimated data symbol
$\tilde{s}$	Final detected data symbol
T	Duration of codeword block with $T = 4N \times T_s$
T_s	Symbol duration
y_r	Channel output signal of received antenna (r)
w_r	Noise signal of receiver antenna (r)
$\hat{w}_r$	FFT of the noise w_r
w	Measurement noise vector
R	Data rate
R_w	Covariance matrix of measurement noise
η	Process noise vector
$z_r^{(v)}$	Output of OFDM demodulator (r) at instant (v)
σ_n^2	Variance of noise w_r
β	Transition coefficient
$\theta_{j,j+1}$	Phase factor $j = 1, 3, \dots, 2N - 1$
φ_ℓ	Phase belonging to a pre-determined set of m^2 different phases
$(\cdot)^H$	Conjugate transpose
$(\cdot)^T$	Transpose
$(\cdot)^*$	Complex conjugate
$E\{\cdot\}$	Expectation
$\star$	Convolution operation
$\ \cdot\ _F$	Frobenius norm
$\otimes$	Kronecker product

MIMO-OFDM channel is estimated by a noise subspace method, exploiting the null space induced by the OSTBC. The work in [13] presented an EVD decomposition of the signal covariance matrices from a group of carefully selected subcarriers to identify the channel with a low order signal constellation.

All these algorithms need a long sequence of data for the receiver to converge, which leads to high complexity. In this paper, a blind method is proposed, based on a new OSTBC which has a very low complexity and which does not need a long sequence to obtain an optimized estimation of the channel. This orthogonal code allows to estimate blind MIMO channels and reduce the complexity of the detector.

The remainder of the paper is organized as follows: in Section 2, a model of the MIMO-OFDM system with the proposed OSTBC is described. Section 3 presents the proposed blind estimation method tracking the time variation of a MIMO-OFDM channel. Section 4 is devoted to a detailed description of the data detector. Simulation results are presented and discussed in Section 5. Finally, some conclusions are drawn in Section 6.

2. MIMO-OFDM System Model

The proposed OSTBC is based on a MIMO-OFDM system model with $N_t = 2$ transmitting antennas, $N_r = 2$ receiving antennas and N subcarriers, as shown in Fig. 1. It is assumed that the channel order does not exceed the cyclic prefix length and remains constant over 4 successive OFDM symbols (codeword length equal to $4N$). In the blind environment, where no prior knowledge about channel information state is available at the receiver, an initial blind channel estimator is used to initiate the steady state Kalman channel estimator (SS-KCE) that refines estimation of the time changing coefficients of the matrix channel, and tracks them. Combining OSTBC and a MIMO-OFDM system can significantly simplify the structure of the receiver and yields a maximum transmit diversity.

The output of the MIMO-OFDM channel from Fig. 1 is expressed as:

$$\begin{aligned} y_1[n] &= \mathbf{h}_{11} \star x_1[n] + \mathbf{h}_{12} \star x_2[n] + w_1[n] \\ y_2[n] &= \mathbf{h}_{21} \star x_1[n] + \mathbf{h}_{22} \star x_2[n] + w_2[n] \end{aligned} \quad (1)$$

where $y_r[n]$ ($r = 1, 2$) is the channel output at time instant $[n]$, $w_r[n]$ ($r = 1, 2$) is the noise at time instant $[n]$, assumed to be a zero mean additive white Gaussian Noise (AWGN) with variance σ_n^2 , $\mathbf{h}_r = [h_r^{(0)}, \dots, h_r^{(L-1)}]$ ($r = 1, 2$ and $t = 1, 2$) is the L length frequency selective sub-channel between the t -th transmitting antenna and the r -th receiving antenna.

It is assumed that the channel matrix $\mathbf{H}$, as defined in Eq. (2), is that of a frequency Rayleigh selective fading MIMO channel that is constant or presents no significant change during the transmitted frame of $4N$ instants.

$$\mathbf{H} = \begin{pmatrix} \mathbf{h}_{11} & \mathbf{h}_{12} \\ \mathbf{h}_{21} & \mathbf{h}_{22} \end{pmatrix}. \quad (2)$$

The transmitted signal $x_t[n]$, from the t -th ($t = 1, 2$) transmitting antenna at time instant $[n]$ is:

$$x_t[n] = \text{Re} \left(e^{2j\pi f_0 n T_s} \sum_{u=1}^{N+N_g} I_t^+[u] \cdot g[n-u] T_s \right), \quad t = 1, 2, \quad (3)$$

where $I_t^+[u]$ is the u -th element of the vector I_t^+ , formed by the output vector from the N length IFFT block, to which is appended a cyclic prefix (CP) of length N_g . Subscript t refers to the t -th antenna.

For more details about the OFDM technique, see [14]–[20].

2.1. Proposed OSTBC

The existing OSTBC solutions were applied to MIMO systems only in the case of a flat fading channel, to ensure that its orthogonal property is preserved. This property is important for reconstructing the transmitted data and reducing complexity of the detector.

In this paper, we propose an OSTBC approach that works even for a frequency selective channel. This code is developed for a $(N_t, N_r) = (2, 2)$ MIMO system and an OFDM system of N subcarriers, as shown in Fig. 2.

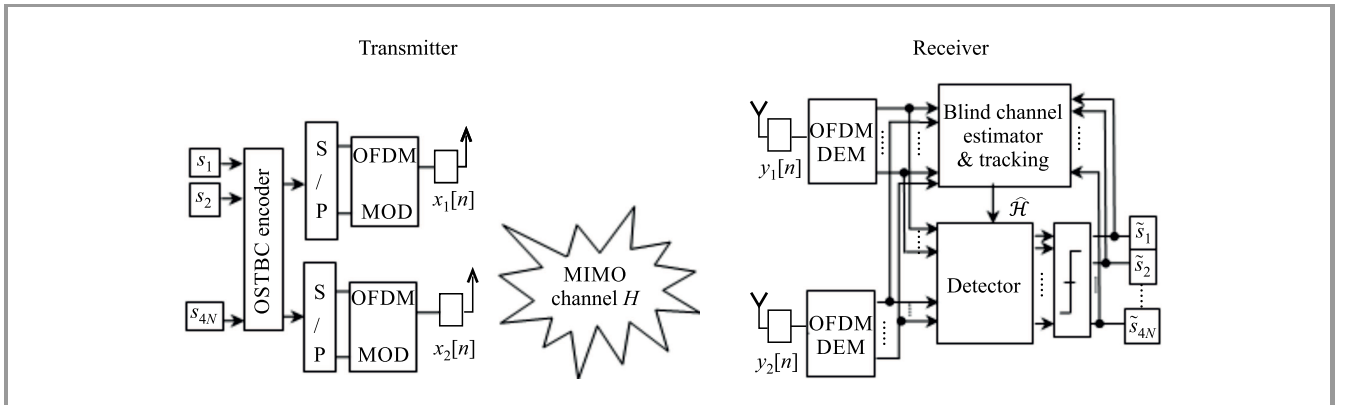


Fig. 1. Block diagram for blind estimation and detection of a MIMO-OFDM system based on the proposed OSTBC.

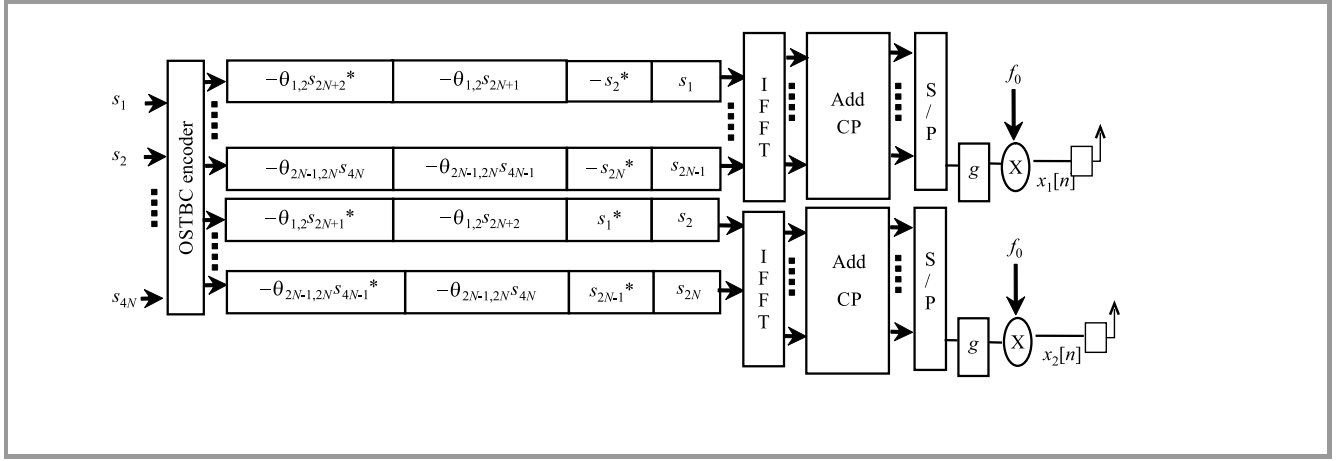
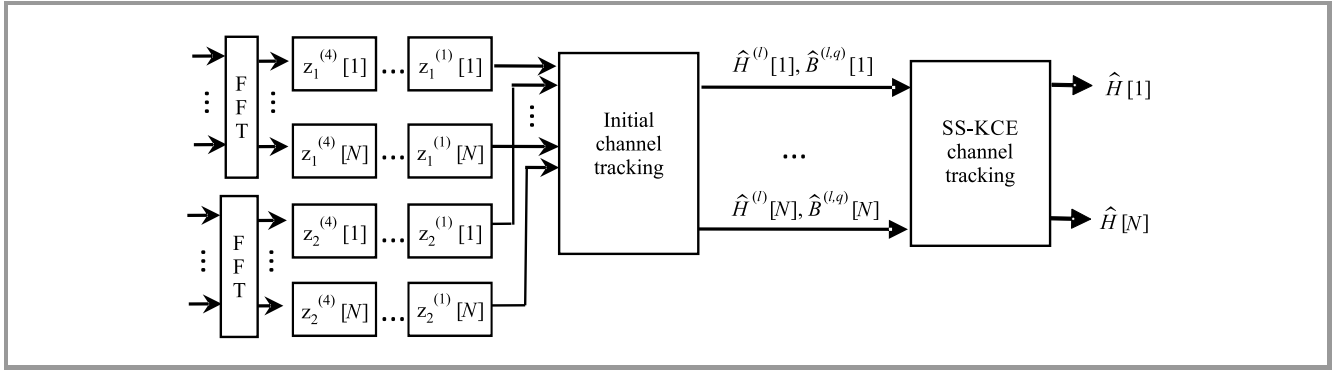

 Fig. 2. OSTBC encoder with OFDM system with N sub-carriers.


Fig. 3. Estimation and tracking of MIMO-OFDM channel.

The generator matrix G of the proposed OSTBC is given by:

$$G = \begin{bmatrix} s_1 & s_2 \\ \vdots & \vdots \\ s_{2N-1} & s_{2N} \\ -s_2^* & s_1^* \\ \vdots & \vdots \\ -s_{2N}^* & s_{2N-1}^* \\ \vdots & \vdots \\ \theta_{2N-1,2N} s_{4N-1} & \theta_{2N-1,2N} s_{4N} \\ -\theta_{1,2} s_{2N+2}^* & \theta_{1,2} s_{2N+1}^* \\ \vdots & \vdots \\ -\theta_{2N-1,2N} s_{4N}^* & \theta_{2N-1,2N} s_{4N-1}^* \end{bmatrix}_{4N \times 2}, \quad (4)$$

where s_i ($i = 1, 2, \dots, 4N$) is the transmitted symbol that belongs to a constellation of m points, $\theta_{j,j+1} = e^{j\varphi_l}$, $j = 1, 3, \dots, 2N-1$ and $\ell \in \{1, 2, \dots, m^2\}$ is a phase factor, with φ_l being a phase belonging to a pre-determined set of m^2 different phases.

Each possible combination of $\begin{bmatrix} s_j \\ s_{j+1} \end{bmatrix}$ has a unique phase factor $\theta_{j,j+1}$. The choice of the pre-determined set of m^2 different phases should respect the following condition:

$$\varphi_{\ell_1} \neq \varphi_{\ell_2} + n\pi \quad \forall \ell_1 \neq \ell_2, \quad \{\ell_1, \ell_2\} \in \{1, 2, \dots, m^2\}. \quad (5)$$

An adequate choice of the phase factor can solve the ambiguity problem during the channel estimation step.

The generator matrix G is composed of four parts. The first two parts (column-wise) are used by the receiver to find all possible combinations for channel matrix estimation. The remaining parts (with the phase factors) are used to eliminate ambiguity in finding the matrix channel when the minimum square error (MSE) criterion is applied. This code is orthogonal, since we have:

$$G^H \times G = \sum_{i=1}^{4N} \|s_i\|^2 I_2. \quad (6)$$

3. Blind Channel Estimation and Tracking

The proposed blind estimator of the MIMO-OFDM channel comprises two parts: initial blind channel estimation algorithm and decoding. It works only for the first received codeword block to initialize the SS-KCE algorithm that tracks the fast channel's variations for the remaining codeword stream.

After the OFDM demodulator, which includes S/P, CP removal, and FFT (see Fig. 3), the signal $z_r^{(v)}$ from the r -th receiving antenna at instant (v) may be expressed as [21]:

$$z_r^{(v)}[u] = s_1^{(v)}[u]h_{r1}[u] + s_2^{(v)}[u]h_{r2}[u] + w_r^{(v)}[u];$$

$$r = 1, 2; u = 1, 2, \dots, N; v = 1, 2, \dots, 4, \quad (7)$$

where $h_{r1}[u]$, $h_{r2}[u]$ are the u -th coefficients of the FFT of the sub-channels $\mathbf{h}_{r1}$ and $\mathbf{h}_{r2}$, respectively, $w_r[u]$ is the u -th coefficient of the FFT of the noise $w_r[n]$, and $s_1^{(v)}[u]$ and $s_2^{(v)}[u]$ correspond to the upper and lower coded data from the OSTBC encoder.

3.1. Initial Blind Estimation of the MIMO-OFDM Channel

Using the first data block received, the blind algorithm based on MSE criteria provides an initial estimation of the channel's FFT coefficient matrices, together with the most probable transmitted codeword block. The two outputs of this algorithm are used to initialize SS-KCE that tracks variations of the channel's coefficients.

$$\begin{aligned} & s_1^{(1)}[u]^* \times z_1^{(1)}[u] - s_2^{(1)}[u] \times z_1^{(2)}[u] \\ &= h_{11}[u] \left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right) \\ &+ s_1^{(1)}[u]^* \times w_1^{(1)}[u] - s_2^{(1)}[u] \times w_1^{(2)}[u]; \\ & s_2^{(1)}[u]^* \times z_2^{(1)}[u] + s_1^{(1)}[u] \times z_1^{(2)}[u] \\ &= h_{12}[u] \left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right) \\ &+ s_2^{(1)}[u]^* \times w_1^{(1)}[u] + s_1^{(1)}[u] \times w_1^{(2)}[u]; \\ & s_1^{(1)}[u]^* \times z_2^{(1)}[u] - s_2^{(1)}[u] \times z_2^{(2)}[u] \\ &= h_{21}[u] \left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right) \\ &+ s_1^{(1)}[u]^* \times w_2^{(1)}[u] - s_2^{(1)}[u] \times w_2^{(2)}[u]; \\ & s_2^{(1)}[u]^* \times z_2^{(1)}[u] + s_1^{(1)}[u] \times z_2^{(2)}[u] \\ &= h_{22}[u] \left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right) \\ &+ s_2^{(1)}[u]^* \times w_2^{(1)}[u] + s_1^{(1)}[u] \times w_2^{(2)}[u]; \end{aligned} \quad (8)$$

By neglecting the noise effect, an estimate of the matrix channel $\hat{\mathcal{H}}[u]$ may be obtained by:

$$\begin{aligned} \hat{h}_{11}[u] &= \frac{s_1^{(1)}[u]^* \times z_1^{(1)}[u] - s_2^{(1)}[u] \times z_1^{(2)}[u]}{\left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right)}; \\ \hat{h}_{12}[u] &= \frac{s_2^{(1)}[u]^* \times z_1^{(1)}[u] + s_1^{(1)}[u] \times z_1^{(2)}[u]}{\left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right)}; \\ \hat{h}_{21}[u] &= \frac{s_1^{(1)}[u]^* \times z_2^{(1)}[u] - s_2^{(1)}[u] \times z_2^{(2)}[u]}{\left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right)}; \\ \hat{h}_{22}[u] &= \frac{s_2^{(1)}[u]^* \times z_2^{(1)}[u] + s_1^{(1)}[u] \times z_2^{(2)}[u]}{\left(\|s_1^{(1)}[u]\|^2 + \|s_2^{(1)}[u]\|^2 \right)}; \end{aligned} \quad (9)$$

Since $s_1^{(1)}[u]$ and $s_2^{(1)}[u]$ are unknown, there are m^2 possibilities for $\hat{\mathcal{H}}[u]$:

$$\hat{\mathcal{H}}^{(\ell)}[u] = \begin{bmatrix} \hat{h}_{11}^{(\ell)}[u] & \hat{h}_{12}^{(\ell)}[u] \\ \hat{h}_{21}^{(\ell)}[u] & \hat{h}_{22}^{(\ell)}[u] \end{bmatrix}, \quad (10)$$

$$\begin{aligned} \hat{h}_{11}^{(\ell)}[u] &= \frac{s_{\ell_1}^* z_1^{(1)}[u] - s_{\ell_2} z_1^{(2)}[u]}{\left(\|s_{\ell_1}\|^2 + \|s_{\ell_2}\|^2 \right)}; \\ \hat{h}_{12}^{(\ell)}[u] &= \frac{s_{\ell_2}^* z_1^{(1)}[u] + s_{\ell_1} z_1^{(2)}[u]}{\left(\|s_{\ell_1}\|^2 + \|s_{\ell_2}\|^2 \right)}; \\ \hat{h}_{21}^{(\ell)}[u] &= \frac{s_{\ell_1}^* z_2^{(1)}[u] - s_{\ell_2} z_2^{(2)}[u]}{\left(\|s_{\ell_1}\|^2 + \|s_{\ell_2}\|^2 \right)}; \\ \hat{h}_{22}^{(\ell)}[u] &= \frac{s_{\ell_2}^* z_2^{(1)}[u] + s_{\ell_1} z_2^{(2)}[u]}{\left(\|s_{\ell_1}\|^2 + \|s_{\ell_2}\|^2 \right)}; \end{aligned} \quad (11)$$

with $\ell = 1, \dots, m^2$, $\{s_{\ell_1}, s_{\ell_2}\} \in \{C\}$, and $u = 1, 2, \dots, N$.

To obtain the right estimate of the channel matrix and, at the same time, detect the four symbols $\{s_{\ell_1}, s_{\ell_2}, s_{q_1}, s_{q_2}\}$, of the transmitted codeword, we apply the following MSE criterion:

$$\begin{aligned} & \langle \hat{\mathcal{H}}^{(\ell)}[u], \hat{\mathbf{B}}^{(\ell,q)} \rangle \geq \argmin_{(\hat{\mathcal{H}}^{(\ell)}[u], \hat{\mathbf{B}}^{(\ell,q)})} \|Z[u] \\ & - \hat{\mathcal{H}}^{(\ell)}[u] \hat{\mathbf{B}}^{(\ell)}[u] \hat{\mathbf{B}}^{(\ell,q)}\|_F^2; \\ Z[u] &= \begin{bmatrix} Z_1^{(1)}[u] & Z_1^{(2)}[u] & Z_1^{(3)}[u] & Z_1^{(4)}[u] \\ Z_2^{(1)}[u] & Z_2^{(2)}[u] & Z_2^{(3)}[u] & Z_2^{(4)}[u] \end{bmatrix}; \\ \hat{\mathbf{B}}^{(\ell,q)} &= \begin{bmatrix} s_{\ell_1} & -s_{\ell_1}^* & \theta_{\ell} s_{q_1} & -\theta_{\ell} s_{q_2}^* \\ s_{\ell_2} & s_{\ell_2}^* & \theta_{\ell} s_{q_2} & \theta_{\ell} s_{q_1}^* \end{bmatrix}; \end{aligned} \quad (12)$$

$\{s_{\ell_1}, s_{\ell_2}, s_{q_1}, s_{q_2}\} \in \{C\}$; $u = 1, 2, \dots, N$; $\ell, q = 1, \dots, m^2$.

The value of θ_{ℓ} is related to the pair of symbols $\begin{bmatrix} s_{\ell_1} \\ s_{\ell_2} \end{bmatrix}$ transmitted by the two antennas.

3.2. SS-KCE for MIMO-OFDM Channel Tracking

The SS-KCE was proposed for tracking the variations of a flat fading channel in an OSTBC-MIMO system [22]. It has the advantage of performing nearly as good as the optimal time-varying Kalman estimator with reduced complexity [23], [24].

In this paper, the property of an OFDM system, such as frequency selective MIMO channel which may be modeled by a set of N parallel flat fading channels is exploited to apply the SS-KCE in the proposed OSTBC-MIMO-OFDM approach.

Two formulas are required for channel tracking: the process equation and the measurement equation [23], [25], [26]. The former describes the dynamic behavior of the state variables to be estimated, while the latter captures the relationship between the output system and the state variables.

In the case concerned, Eq. (13) may be considered as the process equation:

$$\hat{h}_{(k)}[u] = F\hat{h}_{(k-1)}[u] + A\eta[u]; \quad u = 1, 2, \dots, N, \quad (13)$$

where:

- $\hat{h}_{(k)}[u] = \text{vec}(\mathcal{H}_{(k)}[u])$, vec being the operator that places the columns of a matrix on top of each other. $\hat{h}_{(k)}[u]$ is assumed to be nearly constant during a frame with the length $4N$ symbols;
- $\eta[u]$ is a vector of length $N_r \times N_t$ containing independent samples of a circularly symmetric zero-mean FFT of Gaussian noise, with the covariance matrix $R_\eta = I_{N_r N_t}$.
- $F = \beta I_{N_r N_t}$ is a transition diagonal matrix with $\beta = J_0(2\pi f_D T_s)$. J_0 is the first kind zero-order Bessel function $f_D T_s$ is the normalized Doppler rate (assumed to be the same for all transmitting-receiving antenna pairs), and T_s is the baud duration [28].

According to the widely used wide-sense stationary uncorrelated scattering (WSSUS) model [27], the FFT channel coefficients are modeled as independent, zero-mean, complex Gaussian random variables with the auto-correlation matrix given by:

$$E[\hat{h}_{(k)}[u] \hat{h}_{(k+1)}[u]^{\mathcal{H}}] \approx F^{\mathcal{H}}; \quad u = 1, 2, \dots, N. \quad (14)$$

From Eq. (13), the rate of the channel variations is fixed by transition matrix F and the amplitude of all $\hat{h}_{(k)}[u]$ is controlled by matrix A . To be sure that the correlation matrix of $\hat{h}_{(k)}[u]$ is unitary, matrix A should be taken as $A = \sqrt{1 - \beta^2} I_{N_r N_t}$ [24].

For instance, if we have a mobile terminal that moves with the velocity of $v_0 = 90$ km/h, carrier frequency of $f_0 = 2.3$ GHz, and data rate of $R = 4N \times 0.2$ Mb/s, the maximum Doppler frequency is $f_D = f_0 v_0 / c$, where c is the speed of light. The normalized Doppler rate is $f_D T \approx 0.001$, $T = 4N \times T_s$ where $R = 1/T_s$ and N is the number of subcarriers. In this case, according to the Kalman filtering terminology, $\hat{h}_{(k)}[u]$ represents the state vector and the FFT of the channel output, $Z[u]$, defined in Eq. (12), represents the measurement. By stacking the columns of matrices $Z[u]$, $\mathcal{H}_{(k)}[u]$ and $W[u]$, the following measurement equation can be obtained from Eq. (7):

$$z[u] = \mathcal{S}[u] \hat{h}_{(k)}[u] + w[u], \quad u = 1, 2, \dots, N, \quad (15)$$

where: $\mathcal{S}[u] = \mathcal{S}[u]^T \otimes I_{N_r}$,

$$\mathcal{S}[u] = \begin{bmatrix} s_j & -s_{j+1}^* & \theta_{j,j+1} \times s_{2N+j} & -\theta_{j,j+1} \times s_{2N+j+2}^* \\ s_{j+1} & s_j^* & \theta_{j,j+1} \times s_{2N+j+1} & \theta_{j,j+1} \times s_{2N+j}^* \end{bmatrix},$$

$$j = 1, 3, \dots, 2N-1, \quad u = 1, 2, \dots, N,$$

$w[u]$ is the measurement noise with a covariance matrix $R_w = \sigma_w^2 I$.

Then, Eqs. (13) and (15) represent the state-space formulation of the problem of estimating a flat varying MIMO channel. Both are linear functions of state vector $\hat{h}_{(k)}[u]$ and noises $w[u]$ and $\eta[u]$ are independent, white and Gaussian, the Kalman filter provides the optimal recursive estimates in the MMSE sense of the channel coefficients [23], [24]. SS-KF proposed by [22] is extended and used in the present work to track channel variations. In this filter, the channel coefficients are updated by:

$$\hat{h}_{k/k}[u] = \beta B_\infty \hat{h}_{k-1/k-1}[u] + \frac{1}{n_s} A_\infty \mathcal{S}[u]^{\mathcal{H}} z[u], \quad u = 1, 2, \dots, N. \quad (16)$$

$$\text{where } A_\infty = P_\infty \left(\frac{\sigma_w^2}{n_s} I_{N_r N_t} + P_\infty \right)^{-1}, \quad B_\infty = I_{N_r N_t} - A_\infty.$$

The steady-state value of the estimation error covariance matrix P_∞ can be obtained by solving the following discrete algebraic Riccati equation (DARE) [23], [24]:

$$P_\infty = \beta^2 P_\infty - \beta^2 P_\infty \left(P_\infty + \frac{\sigma_w^2}{n_s} I_{N_r N_t} \right)^{-1} P_\infty + \sigma_w^2 I_{N_r N_t}, \quad (17)$$

where n_s represents the energy, supposed to be constant for each uncoded data block.

4. Data Detection

Exploiting the orthogonal property of the proposed OSTBC simplifies the detection process [29]. In this case, estimates $\hat{s}_j$ and $\hat{s}_{j+1}$ are decoupled and can be obtained as:

The u -th channel output can be written as:

$$\begin{bmatrix} z_1^{(1)}[u] & z_1^{(2)}[u] \\ z_2^{(1)}[u] & z_2^{(2)}[u] \end{bmatrix} = \begin{bmatrix} h_{11}[u] & h_{12}[u] \\ h_{21}[u] & h_{22}[u] \end{bmatrix} \begin{bmatrix} s_j & -s_{j+1}^* \\ s_{j+1} & s_j^* \end{bmatrix} + \begin{bmatrix} w_1^{(1)}[u] & w_1^{(2)}[u] \\ w_2^{(1)}[u] & w_2^{(2)}[u] \end{bmatrix}, \quad u = 1, 2, \dots, N. \quad (18)$$

Equation (18) may be rewritten as:

$$\begin{bmatrix} z_1^{(1)}[u] \\ z_2^{(1)}[u] \\ z_1^{(2)*}[u] \\ z_2^{(2)*}[u] \end{bmatrix} = \mathcal{H}[u] \times \begin{bmatrix} s_j \\ s_{j+1} \end{bmatrix} + \begin{bmatrix} w_1^{(1)}[u] \\ w_2^{(1)}[u] \\ w_1^{(2)*}[u] \\ w_2^{(2)*}[u] \end{bmatrix}, \quad (19)$$

where $\mathcal{H}[u]$ is an orthogonal matrix, which groups together two effects: channel attenuation and coding. Its transpose conjugate is given by:

$$\mathcal{H}^{\mathcal{H}}[u] = \begin{bmatrix} \hat{h}_{11}^* & \hat{h}_{21}^*[u] & \hat{h}_{12}[u] & \hat{h}_{22}[u] \\ \hat{h}_{12}^* & \hat{h}_{22}^*[u] & -\hat{h}_{11}[u] & -\hat{h}_{21}[u] \end{bmatrix}. \quad (20)$$

The estimate of $\begin{bmatrix} s_j \\ s_{j+1} \end{bmatrix}$ can be obtained from:

$$\begin{bmatrix} \hat{s}_j \\ \hat{s}_{j+1} \end{bmatrix} = \mathcal{H}^{\mathcal{H}}[u] \times \begin{bmatrix} z_1^{(1)}[u] \\ z_2^{(1)}[u] \\ z_1^{(2)*}[u] \\ z_2^{(2)*}[u] \end{bmatrix}. \quad (21)$$

From this estimate, an estimate of the factor $\theta_{j,j+1}[u]$ is deduced and used to get $\begin{bmatrix} \hat{s}_{j+2N} \\ \hat{s}_{j+2N+1} \end{bmatrix}$:

$$\begin{bmatrix} \hat{s}_{j+2N} \\ \hat{s}_{j+2N+1} \end{bmatrix} = [\hat{\theta}_{j,j+1}[u]]^{-1} \mathcal{H}^{\mathcal{H}}[u] \begin{bmatrix} z_1^{(3)}[u] \\ z_2^{(3)}[u] \\ z_1^{(4)*}[u] \\ z_2^{(4)*}[u] \end{bmatrix}. \quad (22)$$

A threshold detector is applied to get an approximation of the transmitted data symbols:

$$\begin{bmatrix} \hat{s}_j \\ \hat{s}_{j+1} \end{bmatrix} = \mathcal{Q}\left(\begin{bmatrix} \hat{s}_j \\ \hat{s}_{j+1} \end{bmatrix}\right); \quad \begin{bmatrix} \hat{s}_{j+2N} \\ \hat{s}_{j+2N+1} \end{bmatrix} = \mathcal{Q}\left(\begin{bmatrix} \hat{s}_{j+2N} \\ \hat{s}_{j+2N+1} \end{bmatrix}\right). \quad (23)$$

5. Simulation Results

Simulation results are presented below to illustrate the performance of the proposed blind channel estimation and data recovery method. These results were obtained using Monte Carlo runs. In each run, we generate a stream composed of 1,000 blocks of $4 \times N$ BPSK information symbols, with $N = 8, 16, 32, 64, 128, 512$ and 1024 subcarriers, using a cyclic prefix of $CP = 25\% N$. This stream was transmitted through a $(2, 2)$ Rayleigh frequency selective fading MIMO channel, with two different sub-channels length values: $L = 3$ and $L = 4$.

A set of phases $\{\phi\} = \{0, \pi/3, \pi/2, 2\pi/3\}$ that verify Eq. (5) is chosen arbitrarily. The associate set of pairs of symbols is $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}$.

Figure 4 shows the accuracy of time channel variation tracking performed by SS-KCE, in the case of a channel speed variation factor of $f_D T = 0.001$, with $T = 4NT_s$, and different values of subcarriers N . It plots the normalized mean square error (NMSE) of the channel estimation, defined in Eq. (22), versus the SNR.

$$\text{NMSE} = \frac{1}{N_{MC}} \sum_{ir=1}^{N_{MC}} \frac{\|\hat{H}_{ir} - H\|_F^2}{\|H\|_F^2}, \quad (24)$$

where N_{MC} represents the number of Monte Carlo runs.

For $N = 16$, Fig. 4 shows that the blind channel estimation is better (small values of NMSE) compared to other values of $N > 16$, so raising the number of subcarriers N may increase the error channel estimation NMSE. This is because of N cumulative errors issued from N banks of the SS-KCE channel estimator. However, performance is not good when N is low, because in the OFDM technique, in order to convert a frequency selective channel to a set of parallel

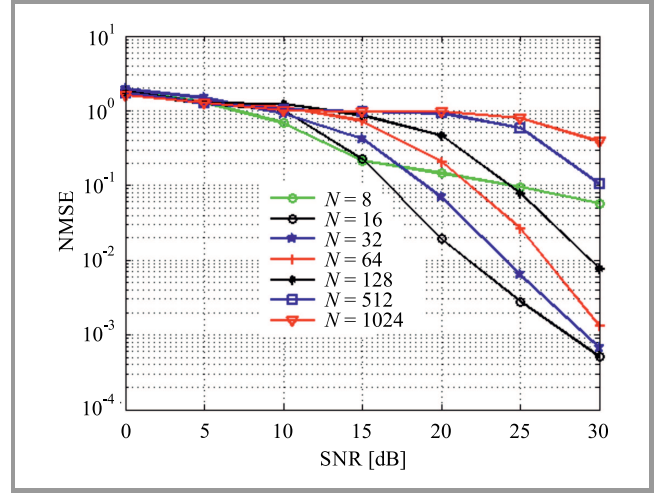


Fig. 4. Channel estimation NMSE of the proposed blind MIMO-OFDM algorithm with $(N_t, N_r) = (2, 2)$, BPSK, $L = 3$, $f_D T = 0.001$ and different numbers of N subcarriers.

flat fading sub-channels, a high value of subdivisions N is required that is not sufficient for $N = 8$. So, in this case, the channel is still not totally flat fading after the conversion and, as a consequence, the error channel estimation NMSE is raised.

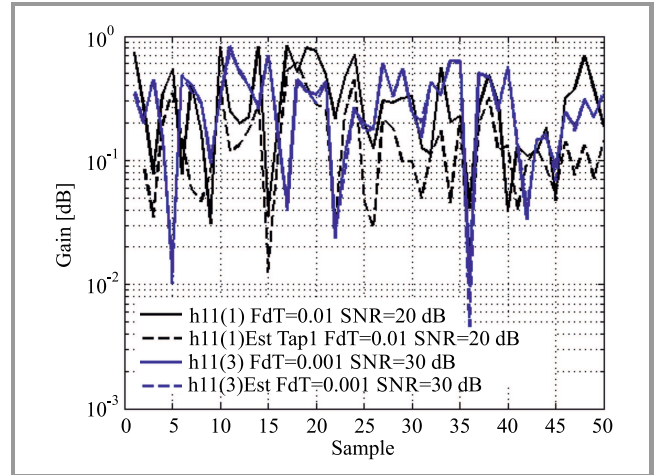


Fig. 5. SS-KCE tracking of channel coefficients h_{11} (1) and h_{11} (3) for $f_D T = 0.01$ at SNR = 20 dB and $f_D T = 0.001$ at SNR = 30 dB, respectively.

To illustrate the time tracking capability of the SS-KCE, coefficients h_{11} (1) and h_{11} (3) of the first MIMO selective sub-channel for $f_D T = 0.01$ at SNR = 20 dB, and $f_D T = 0.001$ at SNR = 30 dB, respectively, are compared with their estimates in Fig. 5. It is clear that the higher the SNR is and the smaller $f_D T$ is, the more accurate the channel estimation performed by the SS-KCE is.

In Fig. 6, the bit error rate (BER) of the proposed algorithm is plotted versus SNR for $f_D T = 0.001$ and different values of the number of subcarriers N . One may observe that when the number of subcarriers N is very high, the performance of the SS-KCE degrades. This is due to the estimation errors of the set of N sub-channel matrices, which increase

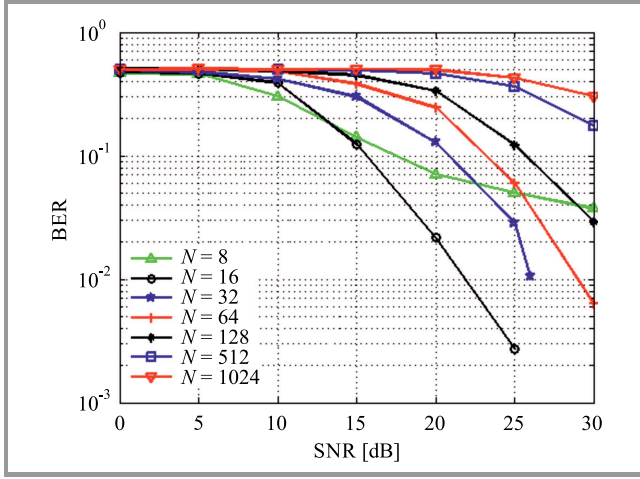


Fig. 6. BER of the proposed blind MIMO-OFDM algorithm with $(N_t, N_r) = (2, 2)$, BPSK, $L = 3$, $f_D T = 0.001$ and different values for the number of subcarriers N .

as N increases. These errors exert a negative impact on the performance of the detector, which in turn degrades the performance of the channels' estimation by SS-KCE. On the one hand, N should be large enough to guarantee that the MIMO selective channel is equivalent to a set of parallel flat fading channels. For the case of $N = 8$, BER is not good, because the detector is based directly on the output of the blind channel estimator, which is not very accurate (NMSE of $N = 8$ – Fig. 4). On the other hand, too large a value of N degrades the performance of the receiver. Therefore, the choice of N must result from a trade-off between these two contradictory requirements.

Finally, in Fig. 7 the performance of the proposed blind receiver with SS-KCE tracking of the channel variations is compared with the performances of two other blind MIMO-OFDM algorithms proposed in [30] and [31], respectively, for: $f_D T = 0$, $N = 64$, $L = 4$, $(N_t, N_r) = (2, 2)$, and BPSK. As one may observe, the proposed algorithm outperforms these two algorithms when SNR is higher than 20 dB.

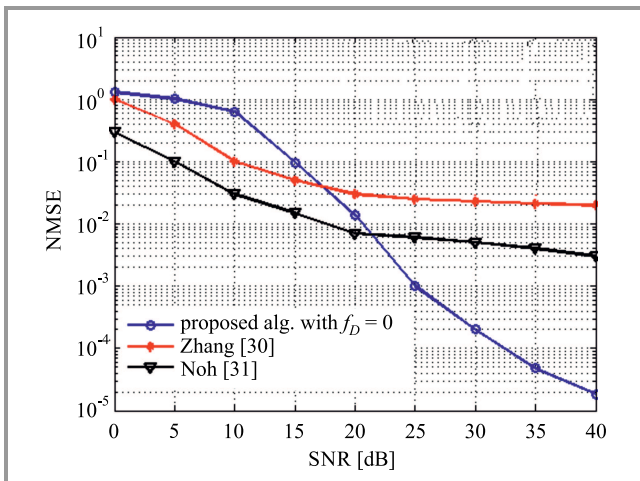


Fig. 7. Comparison between the proposed algorithm and two other blind MIMO-OFDM algorithms for $(N_t, N_r) = (2, 2)$, $N = 64$, BPSK, and $L = 4$.

6. Conclusion

In this paper, a blind MIMO system that operates over a frequency selective fading channel is proposed. It extends the already proposed OSTBC to be used with the well-known OFDM technique to solve the problem of channel frequency selectivity. This OSTBC offers a double benefit: on the one hand, it allows to alleviate the ambiguity that occurs in the estimation of the channel matrix with judicious choice of the number of subcarriers N , and on the other hand, it reduces complexity of the detector thanks to its orthogonal property. The performance of the proposed structure has been evaluated by means of Monte Carlo simulations and compared with two other similar systems proposed in literature. The results obtained show that the proposed system has the best performance for a reasonable range of SNR values.

References

- [1] A. Sibille, C. Oestges, and A. Zanella, *MIMO: From Theory to Implementation*. Academic Press, 2010 (ISBN: 9780123821959).
- [2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE J. on Sel. Areas in Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998 (doi: 10.1109/49.730453).
- [3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction", *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1998 (doi: 10.1109/18.661517).
- [4] F. Boccardi *et al.*, "Multiple antenna techniques in LTE-advanced", *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 114–121, 2012 (doi: 10.1109/MCOM.2012.6163590).
- [5] Y. Aimer, B. S. Bouazza, S. Bachir, and C. Duvanaud, "Interleaving technique implementation to reduce PAPR of OFDM Signal in presence of nonlinear amplification with memory effects", *J. of Telecommun. and Inform. Technol. (JTIT)*, no. 3, pp. 14–22, 2018 (doi: 10.26636/jtit.2018.123517).
- [6] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*. Springer, 2019 (ISBN: 978-3-319-92089-4).
- [7] M. Jiang and L. Hanzo, "Multi-user MIMO-OFDM for next generation wireless systems", *Proc. of the IEEE*, vol. 95, no. 7, pp. 1430–1469, 2007 (doi: 10.1109/JPROC.2007.898869).
- [8] J. D. Jiang, T. C. Lin, and S. M. Phoong, "New subspace-based blind channel estimation for orthogonally coded MIMO-OFDM systems", in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Process. ICASSP 2014*, Florence, Italy, 2014, pp. 6489–6493 (doi: 10.1109/ICASSP.2014.6854854).
- [9] C.-C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging", *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1539–1544, 2010 (doi: 10.1109/TVT.2009.2039008).
- [10] N. Sarmadi, S. Shahbazpanahi, and A. B. Gershman, "Blind channel estimation in orthogonally coded MIMO-OFDM systems: A semidefinite relaxation approach", *IEEE Trans. on Sig. Process.*, vol. 57, no. 6, pp. 2354–2364, 2009 (doi: 10.1109/TSP.2009.2016887).
- [11] N. Sarmadi and M. Pesavento, "Closed-form blind channel estimation in orthogonally coded MIMO-OFDM systems: A simple strategy to resolve non-scalar ambiguities", in *Proc. IEEE 12th Int. Worksh. on Sig. Process. Adv. in Wirel. Commun.*, San Francisco, CA, USA, 2011, pp. 321–325 (doi: 10.1109/SPAWC.2011.5990422).
- [12] B. Zhang, J. L. Yu, and J. W. Lai, "A fast subspace channel estimation for STBC-based MIMO-OFDM systems", in *Proc. 11th Int. Symp. on Wirel. Commun. Sys. ISWCS 2014*, Barcelona, Spain, 2014, pp. 75–79 (doi: 10.1109/ISWCS.2014.693323).

- [13] W. Zhang, F. Gao, and Q. Yin, "Blind channel estimation for MIMO-OFDM systems with low order signal constellation", *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 499–502, 2015 (doi: 10.1109/CLOMM.2015.2393889).
- [14] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission", *The Bell System Tech. J.*, vol. 45, no. 10, pp. 1775–1796, 1966 (doi: 10.1002/j.1538-7305.1966.tb02435.x).
- [15] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform", *IEEE Trans. on Commun.*, vol. 19, no. 5, pp. 628–634, 1971 (doi: 10.1109/TCOM.1971.1090705).
- [16] R. G. Vaughan, "Polarization diversity in mobile communications", *IEEE Trans. on Veh. Technol.*, vol. 39, no. 3, pp. 177–186, 1990 (doi: 10.1109/25.130998).
- [17] R. Van Née and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House Publ., 2000 (ISBN: 978-0890065303).
- [18] J. J. van de Beek, P. Ödling, S. K. Wilson, and P. O. Börjesson, "Orthogonal frequency division multiplexing (OFDM)" in *Review of Radio Science 1996-1999. Comm. C: Signals and Systems*, E. Bonek and W. R. Stone, Eds. International Union of Radio Science (URSI), 1999 (ISBN: 0198565712).
- [19] P. Uthansakul and M. E. Bialkowski, "Multipath signal effect on the capacity of MIMO, MIMO-OFDM and spread MIMO-OFDM", in *Proc. 15th Int. Conf. on Microw., Radar and Wirel. Commun. MIKON-2004*, Warsaw, Poland, 2004, vol. 3, pp. 989–992 (doi: 10.1109/MIKON.2004.1358536).
- [20] H. Yang, "A road to future broadband wireless access: MIMO-OFDM-based air interface", *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 53–60, 2005 (doi: 10.1109/MCOM.2005.1381875).
- [21] I. Barhum, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels", *IEEE Trans. Sig. Process.*, vol. 51, no. 6, pp. 1615–1624, 2001 (doi: 10.1109/TSP.2003.811243).
- [22] M. B. Loiola, R. R. Lopes, and J. Romano, "Kalman filter-based channel tracking in MIMOOSTBC systems", in *Proc. IEEE Global Telecommun. Conf. GLOBECOM 2009*, Honolulu, HI, USA, 2009, pp. 1–4 (doi: 10.1109/GLOCOM.2009.5425545).
- [23] D. Simon, *Optimal State Estimation Kalman, H Infinity, and Non-linear Approaches*. Wiley, 2006 (ISBN: 9780471708582).
- [24] M. B. Loiola and R. R. Lopes, "A state-space approach to semi-blind signal detection in fast frequency-selective fading MIMO channels", in *Proc. IEEE 9th Worksh. on Sig. Process. Adv. in Wirel. Commun. SPAWC 2008*, Recife, Brazil, 2008, pp. 276–280 (doi: 10.1109/SPAWC.2008.4641613).
- [25] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000 (ISBN: 978-0130224644).
- [26] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice Hall, 2002 (ISBN: 9780130901262).
- [27] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*, 2nd ed. Kluwer, 2000 (ISBN: 978-0306462672).
- [28] B. Chen and A. P. Petropulu, "Frequency domain blind MIMO system identification based on second and higher order statistics", *IEEE Trans. Sig. Process.*, vol. 49, no. 8, pp. 1677–1688, 2001 (doi: 10.1109/78.934137).
- [29] H. Jafarkhani, *Space-Time Coding*. Cambridge, U.K.: Cambridge University Press, 2005 (ISBN: 978-0521842914).
- [30] R. Zhang, "Blind channel estimation for precoded OFDM system", in *Proc. of the IEEE Int. Conf. Acoust., Speech, and Sig. Process. ICASSP'05*, Philadelphia, PA, USA, 2005 (doi: 10.1109/ICASSP.2005.1415748).
- [31] S. Noh and M. D. Zoltowski, "A new precoder design for blind channel estimation in MIMO-OFDM systems", in *Proc. IEEE Global Commun. Conf. GLOBECOM 2013*, Atlanta, GA, USA, 2013, pp. 3306–3311 (doi: 10.1109/GLOCOM.2013.6831582).



Abdelkader Tami received his B.Eng. in 2000 from the Telecommunications Institute, Oran, Algeria, M.Sc. in Telecommunications Systems and Micro-waves from the University of Saida, Algeria in 2007, and Ph.D. from University of Science and Technology of Oran, Algeria, in 2016. He is an Associate Professor at the Department of Electronics at Dr. Moulay Tahar University of Saida, Algeria. He is a membership of the Laboratory of Technology of Communications (LTC). His research interests are in MIMO wireless systems and advanced digital communication.

 <https://orcid.org/0000-0001-6041-5521>
 E-mail: abdelkader.tami@hotmail.com
 Laboratory Technology of Communication
 Faculty of Technology
 University of Saida
 Algeria



Mokhtar Keche received his B.Eng. degree in Telecommunications from Ecole Nationale Supérieure des Telecommunications (ENST), Paris, France in 1978, M.Sc. and Ph.D. degrees from the University of Rennes in France and the University of Nottingham, UK, in 1982 and 1998, respectively. He is a Professor at the Department of Electronics, University of Sciences and Technology, Mohamed Boudiaf, Oran, Algeria. His research interests are in the area of digital communications, array processing and multi-target tracking.

E-mail: m.keche@yahoo.com
 Laboratory of Signal and Image
 Department of Electronics University of Sciences and Technology
 Oran, Algeria



Boubaker Seddik Bouazza is an Associate Professor at Dr. Moulay Tahar University, Saida, Algeria. He is the Head of the Laboratory of Communication Technology (LTC). His research work focuses mainly on channel coding for communication systems.

E-mail: bsbouazza@yahoo.fr
 Laboratory of Communication Technology
 Faculty of Technology
 University of Saida
 Algeria

Complex Signal Mapping for Improving Spectral Efficiency of Li-Fi Systems

Ibrahim K. Sileh, Ayad A. Abdulkafi, Mohammed K. Hussein, and Saad M. Hardan

Electrical Engineering, Tikrit University, Salahaddin, Iraq

<https://doi.org/10.26636/jtit.2019.131719>

Abstract—Light fidelity (Li-Fi) systems based on the orthogonal frequency division multiplexing (OFDM) scheme have gained more attention in the communications circles, as a means to provide high spectral efficiency and ensure stricter data rate requirements for visible light communication (VLC) systems. However, common OFDM schemes employ Hermitian symmetry to obtain a real-valued signal which is necessary in intensity modulation and direct detection (IM/DD) optical systems, at the expense of doubling the required bandwidth. In this paper, a novel transmission scheme for an OFDM-based Li-Fi system is proposed to tackle the issue in question. A new approach to complex signal mapping (CSM), based on the pairing function method, has been developed for Li-Fi systems. It does not require Hermitian symmetry and, hence, saves about 50% of the required bandwidth. Unlike existing OFDM-based VLC approaches, the proposed scheme employs CSM to ensure a real and positive signal without Hermitian symmetry in order to fully utilize the bandwidth available to Li-Fi networks. Simulation results show that the proposed scheme significantly outperforms other systems in terms of spectral efficiency. The CSM-OFDM based Li-Fi system also achieves a good peak-to-average power ratio (PAPR) reduction with acceptable bit-error-rate (BER) performance, compared to conventional approaches.

Keywords—Li-Fi, OFDM, signal mapping, PAPR, BER.

1. Introduction

Light fidelity (Li-Fi) technology is an emerging sector on the optical communications arena. It relies on visible light communication (VLC) systems using light generated by light-emitting diodes (LEDs) to transmit data wirelessly. This technology may be useful in supporting and complementing the existing Wi-Fi systems, since its spectrum bandwidth is much broader. Higher data rates and capacities, accommodating more users, can be provided as well. Report [1] states that the Li-Fi market is predicted to rise to reach the value of USD 75.5 billion by 2023, with an annual growth rate of 77%. Along with that, ITU [2] and IEEE [3] have established first generation standards for VLC systems. Several vendors, such as Philips, PureLiFi, etc., have already produced first Li-Fi modems as well.

As LEDs have a signaling bandwidth of limited to several megahertz, orthogonal frequency division multiplexing (OFDM) may be considered a suitable candidate for LED-based Li-Fi systems due to its robustness to time-varying multipaths, its capacity to combat inter-symbol interferences (ISI), its high spectral efficiency and simple equalization [4]. Since the modulating signals of LED have to be real-valued and non-negative in intensity modulation and direct detection (IM/DD) systems, OFDM has to be modified accordingly. The most popular formats are DC-biased optical OFDM (DCO-OFDM) [5], and asymmetrically clipped OFDM (ACO-OFDM) [4]. In these forms, the input data vector of inverse fast Fourier transform (IFFT) has to be constrained to have Hermitian symmetry by adding the conjugate of the complex signal before the IFFT [4]. This ensures a real-valued time domain signal, but also reduces spectral efficiency and doubles the required bandwidth.

In the DCO-OFDM case, the non-negative signal is ensured after adding a DC-bias to the signal. However, due to the high peak-to-average power ratio (PAPR) of OFDM, some negative peaks of the signal will be clipped and the resulting distortion causes bit-error-rate (BER) performance degradation, even with a high bias [6]. ACO-OFDM converts the bipolar signal to its unipolar counterpart by clipping all negative values at zero and then modulating odd subcarriers only. However, a large part of bandwidth has been sacrificed to guarantee the asymmetrical property. Due to this, the spectral efficiency of DCO-OFDM is about twice as large as the spectral efficiency of ACO-OFDM for the same order of quadrature amplitude modulation (M-QAM). Instead of applying Hermitian symmetry and signal clipping, a simple idea has been presented in [7] to ensure real-valued OFDM signals by mapping the complex time domain (output of IFFT) into a real value, with the same symbols prior to QAM modulation in the frequency domain. The work in [8] presented a new coding scheme by employing zero cross correlation (ZCC) codes to achieve the same goal and to eliminate multiple access interference, but this required large set sizes and, hence, more bandwidth. Motivated by this and in order to fully utilize the available bandwidth, this paper proposes a new approach based on complex signal mapping (CSM) with the pairing function method to

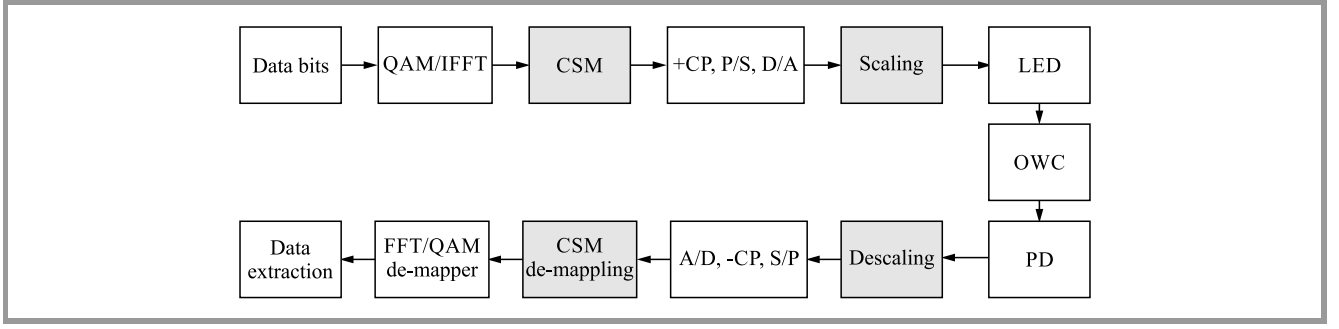


Fig. 1. System block of a Li-Fi transceiver based on OFDM with CSM.

encode a complex OFDM signal into a real-positive signal. This will result in improving the performance of Li-Fi systems utilizing OFDM, in terms of bandwidth efficiency, BER performance and PAPR reduction.

This paper is organized as follows: in Section 2, system descriptions are explained. OFDM CSM procedures are given in Section 3. Detailed simulation results and a discussion are presented in Section 4, while our concluding remarks are provided in Section 5.

2. System Description

Figure 1 represents the proposed CSM-OFDM-based Li-Fi transceiver. The serial data bits are grouped into symbols and modulated using a QAM modulator. The modulated symbols are assigned to sub-carriers and then feed to an N-points IFFT to provide the OFDM time domain signal. The simulation model is similar to the conventional optical OFDM-based VLC, only Hermitian symmetry is not needed for the CSM-OFDM system, and the real-valued time domain signal can be obtained using the CSM that maps the time domain complex OFDM symbols into real integer symbols. A fixed length cyclic prefix (CP) is added to avoid any ISI, followed by parallel-to-serial (P/S) and digital-to-analog (D/A) conversions. The resulting signals are then scaled to be compatible with the LED dynamic range and pass through the optical wireless channel (OWC) with an additive white Gaussian noise (AWGN). On the receiver side, as shown in Fig. 1, data retrieval is the reverse of the process adopted on the transmitter side.

3. Complex Signal Mapping Mechanism

The general mapping operation for a complex signal may be expressed as [7]:

$$x_R = f(x_C) = f(s, a, b), \quad (1)$$

where x_C is the complex signal output of IFFT, which may be written as $a + ib$ in which a is the real part and b is the imaginary part, while s represents all possible signs to each pair of real and imaginary samples, and it may be mapped according to Table 1.

Table 1
Sign mapping

Possible sign pair	Binary representation	Decimal representation, s
--	00	0
-+	01	1
+ -	10	2
++	11	3

Function f maps the complex signal x_C into real-valued x_R with the help of the pairing function at the transmitter, and the estimated $\hat{x}_C$ can be extracted at receiver by de-mapping $\hat{x}_R$ using the inverse of the pairing function. However, the pairing function deals with non-negative integers [9]. Therefore, fractional a and b need to be modified and rounded to integers after taking their absolute values and multiplying them by a certain precision P :

$$A = \lfloor |a| \times P \rfloor, \quad \lfloor |b| \times P \rfloor, \quad (2)$$

where $\lfloor \cdot \rfloor$ is the round function which returns the closest integer value of the given non-integer.

Using the Cantor pairing function, the pair $[A, B]$ can be paired (encoded) into a single non-negative integer C that is uniquely associated with that pair [9]. The pair function is given by:

$$C = \text{pair}[A, B] = \frac{A^2 + 3A + 2AB + B + B^2}{2}. \quad (3)$$

The output of CSM is then to append the sign code with C , i.e. $x_{CSM} = [s : C]$. An example of the complete output code-word is shown in Table 2 for some selected IFFT outputs with P of 10^4 .

The inverse function of unpair outputs the pair associated with each non-negative integer C :

$$A = C - \frac{w^2 + w}{2} \quad \text{and} \quad B = \frac{w^2 + 3w}{2} - C. \quad (4)$$

With intermediate w equaling:

$$w = \left\lfloor \frac{-1 + \sqrt{1 + 8C}}{2} \right\rfloor. \quad (5)$$

Table 2
CSM approach

Selected IFFT output	s	A	B	C	Output codeword $x_{CSM} = [s : C]$
$-0.0738 - i0.1398$	0	738	1398	23145	023145
$-0.0245 + i0.0682$	1	245	682	4346	14346
$0.0321 - i0.0619$	2	321	619	4527	24527
$0.0340 + i0.0849$	3	34	849	7225	37225

If a free noise case is assumed, then the complex signal prior to FFT can be obtained by the reverse process that consists in unparing the received C , dividing A and B by P to extract a and b , and then combining them with their signs to form $a + ib$ as mentioned earlier.

The resulting signal from mapping the complex x_{CSM} signal needs to be scaled according to a specific scaling factor (SF), as shown later. Biasing is then performed by adding a DC-bias to the resulting signal to adjust the minimum value of the scaled signal to turn on voltage (TOV) of the LED. In this way, the maximum value of the signal will be fitted to the saturated point of the LED due to the de-amplification process. The dynamic range of the resulting DRx_{CSM} signal may be defined as:

$$DRx_{CSM} = \max(x_{CSM}) - \min(x_{CSM}) , \quad (6)$$

where $\max(x_{CSM})$ is the maximum value of x_{CSM} and $\min(x_{CSM})$ is its minimum value. LED dynamic range is:

$$DR_{LED} = I_H - I_L , \quad (7)$$

where I_H is the maximum allowed LED drive current and I_L corresponds to the minimum LED drive current according to the LED data sheet. The scaling factor SF can be obtained as:

$$SF = \frac{DR_{LED}}{DRx_{CSM}} . \quad (8)$$

Hence, the de-amplified signal x_D is:

$$x_D = SF \times x_{CSM} , \quad (9)$$

The PAPR is defined as the ratio of the maximum transmitted signal power to the average power:

$$PAPR = 10 \log \frac{\max(x_D^2)}{E[x_D^2]} , \quad (10)$$

where $\max_D^2$ is the maximum value of the de-amplified signal power, $E[\cdot]$ is the average of those values [7]. The complementary cumulative distribution function (CCDF) of PAPR is the most commonly used measure for illustrating the PAPR reduction and it is used here to evaluate PAPR reduction.

The optimum biasing level may be obtained via:

$$DC_{bias} = TOV - \min(x_D) . \quad (11)$$

Finally, the LED input signal can be obtained after shifting above signal to its TOV:

$$x_{LED} = x_D + DC_{bias} = x_D + TOV - \min(x_D) . \quad (12)$$

The spectral efficiency is defined as the amount of information bits transmitted per time unit per available bandwidth [10]. As the bandwidth is fully utilized, the SE of CSM-OFDM-based Li-Fi is twice that of DCO-OFDM, and four times that of ACO-OFDM. For N-FFT points and M QAM order, the SE of Li-Fi system based on CSM-OFDM η_{CSM} is:

$$\eta_{CSM} = 2\eta_{DCO} = 4\eta_{ACO} = \frac{N-2}{N+N_{CP}} \log_2 M , \quad (13)$$

where η_{DCO} [bps/Hz] is the SE of DCO-OFDM system, η_{ACO} is the SE of ACO-OFDM system and N_{CP} is the CP length.

4. Results and Discussion

In this section, simulation results of the proposed scheme, related to spectral efficiency, PAPR and BER performance, are discussed and compared with conventional OFDM approaches in Li-Fi systems. The simulation is carried out using Matlab with 16 QAM and 128 subcarriers. CSM scheme processing is applied to the time domain signal after the IFFT, to obtain a real non-negative signal without employing the Hermitian symmetry to the input signal to fully utilize the available bandwidth in a Li-Fi network, as shown in Fig. 2. It may be noted that there is no distortion as all signals (x_{CSM} , x_D and x_{LED}) have the same waveform. In fact, they are just scaled and shifted.

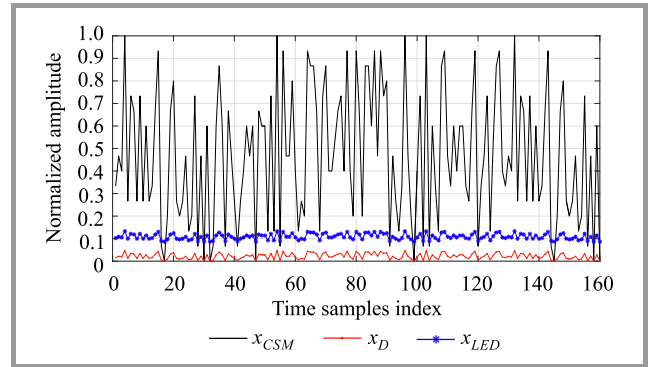


Fig. 2. Time domain signals of an OFDM-based Li-Fi system.

Spectral efficiency performance, as a function of modulation order, for different schemes using 128 FFT/IFFT subcarriers, is illustrated in Fig. 3. The proposed CSM-OFDM avoids the use of Hermitian symmetry, therefore saving bandwidth compared to conventional DCO-OFDM and ACO-OFDM systems. It is clear that the proposed CSM-OFDM scheme outperforms other methods in terms of spectral efficiency for all QAM orders. This is because all available bandwidth has been fully utilized in the case of the proposed CSM-OFDM. DCO-OFDM imposes Hermitian symmetry to generate real-time domain signals and,

hence, spectral efficiency is reduced by half in comparison to CSM-OFDM, while ACO-OFDM uses only one quarter of the available bandwidth. Hence, its spectral efficiency is also one quarter of the proposed CSM-OFDM-based Li-Fi system.

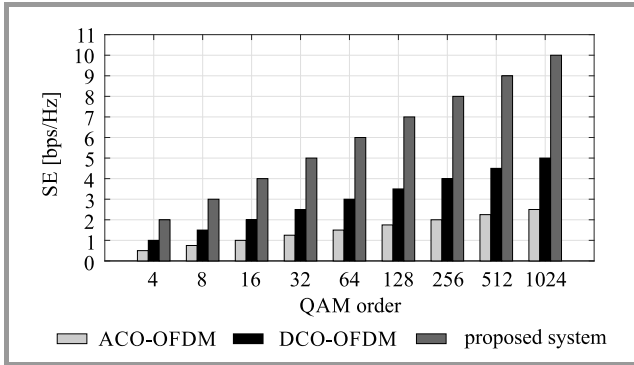


Fig. 3. Spectral efficiency versus modulation order for different schemes.

In order to make a fair comparison between the proposed CSM-OFDM scheme, ACO-OFDM and DCO-OFDM, simulations have been performed with the same size of IFFT and same M-QAM, i.e. IFFT size is 128 and 16-QAM is used for signal modulation for all optical OFDM formats. Figure 4 shows the CCDF comparison of PAPR for CSM-OFDM, ACO-OFDM and DCO-OFDM-based Li-Fi systems. CCDF curves show that DCO-OFDM has a higher PAPR compared to ACO-OFDM and CSM-OFDM systems. Clearly, the CSM-OFDM scheme achieves acceptable PAPR reduction, compared to other methods. However, the proposed scheme has a slightly lower PAPR (0.2 dB gain reduction) compared to ACO-OFDM at around CCDF of 0.1.

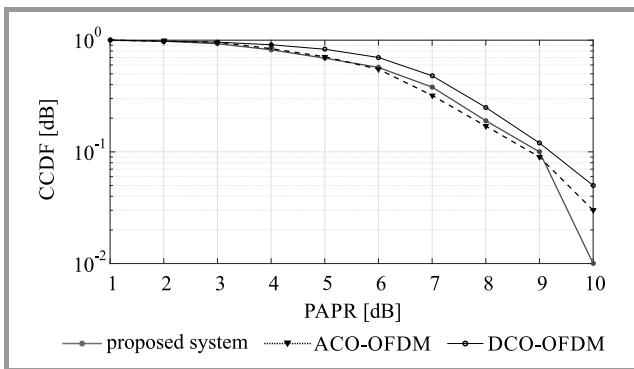


Fig. 4. CCDF performance of the OFDM-based Li-Fi signal with different approaches.

In particular, at $\text{CCDF} = 10^{-1}$, the PAPR threshold values of ACO-OFDM, CSM-OFDM and DCO-OFDM are 8.8, 9 and 9.2 dB, respectively. Figure 5 shows that the proposed scheme also has a good BER performance. It is clear that DCO-OFDM outperforms other systems for all SNR values. The CSM-OFDM achieves better BER than ACO-OFDM for SNR of 19 dB.

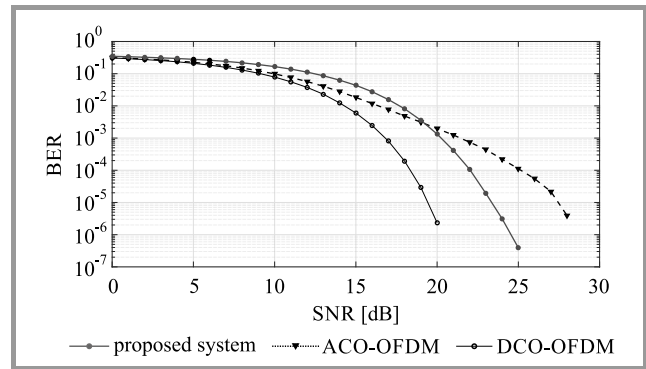


Fig. 5. Comparison of BER performance of OFDM-based Li-Fi signal with different approaches.

In fact, CSM-OFDM is more sensitive to channel change than other two systems at low SNR values, and it is suitable for high SNR when the channel is more slowly time-varying, which is the case in optical wireless communication systems.

5. Conclusion

In this paper, a new approach that avoids the use of Hermitian symmetry and signal clipping for transmitting OFDM signals is developed to increase the bandwidth efficiency of Li-Fi networks. The proposed CSM scheme encodes complex OFDM signals into real and positive values and decodes them prior to FFT at the receiver side, with the help of the pairing function mapping concept. The proposed CSM-OFDM-based Li-Fi system achieves double the spectral efficiency of DCO-OFDM and four times the spectral efficiency of ACO-OFDM systems. Simulation results show that the proposed CSM-OFDM scheme achieves performance that is comparable to ACO-OFDM and DCO-OFDM, with a much lower bandwidth in terms of PAPR reduction and BER. In order to obtain further improvement in spectral efficiency of Li-Fi systems, the concepts of encoding the complex OFDM signal into real and non-negative values for the IM/DD system need to be studied and will be included in future works.

References

- [1] "Li-Fi market size forecast worth \$75.5 billion by 2023" [Online]. Available: <https://www.gminsights.com/pressrelease/LiFi-market>
- [2] "IEEE 802.15 WPAN Task Group 13 (TG13) Multi-Gigabit/s Optical Wireless Communications" [Online]. Available: <http://www.ieee802.org/15/pub/TG13.html>
- [3] "High speed indoor visible light communication transceiver – System architecture, physical layer and data link layer specification", [811-WP1] Draft new Recommendation ITU-T G.vlc [Online]. Available: <https://www.itu.int/md/T13-SG15-160919-TD-WP1-0811/en>
- [4] J. Armstrong and A. Lowery, "Power efficient optical OFDM", *Electron. Lett.*, vol. 42, no. 6, pp. 370–372, 2006 (doi: 10.1049/el:20063636).
- [5] O. Gonzalez, R. Perez-Jimenez, S. Rodriguez, J. Rabadan, and A. Ayala, "Adaptive OFDM system for communications over the indoor wireless optical channel", *IEE Proc. – Optoelec.*, vol. 153, pp. 139–144, 2006 (doi: 10.1049/ip-opt:20050081).

- [6] Y. Aimer, B. S. Bouazza, S. Bachir, and C. Duvanaud, "Interleaving technique implementation to reduce PAPR of OFDM signal in presence of nonlinear amplification with memory effects", *J. of Telecommun. and Inform. Technol.*, no. 3, 2018 (doi: 10.26636/jtit.2018.123517).
- [7] A. A. Abdulkafi, M. Y. Alias, and Y. S. Hussein, "A novel approach for PAPR reduction in OFDM-based visible light communications", in *Proc. Int. Conf. on Plat. Technol. and Service PlatCon 2017*, Busan, South Korea, 2017, pp. 1–4 (doi: 10.1109/PlatCon.2017.7883678).
- [8] M. Addad and A. Djebbari, "A new code family for QS-CDMA visible light communication systems", *J. of Telecommun. and Inform. Technol.*, no. 3, 2018 (doi: 10.26636/jtit.2018.124318).
- [9] S. Wolfram, *A New Kind of Science*. Wolfram Media, 2002 (ISBN13: 9781579550080).
- [10] D. Tsonev, M. S. Islim, and H. Haas, "OFDM-Based visible light communications", in *Optical Wireless Communications. An Emerging Technology*, M. Uysal, C. Capsoni, Z. Ghassemloooy, A. Boucouvalas, and E. Udvarý, Eds. Springer, 2016, pp. 255–298 (doi: 10.1007/978-3-319-30201-0.12)



Ibrahim K. Sileh received his B.Sc. degree in Electrical Engineering (majoring in Electronic and Communications) from Al-Mosul University, Mosul, Iraq, in 1993, and M.Sc. degree in Communication Engineering from University of Technology: Al-Rasheed College, Baghdad, Iraq, in 2001. He received his Ph.D. in Communication Engi-

neering from University of Southern Queensland, Australia in 2014. He is a staff member at the College of Engineering, Tikrit University, Iraq. His research interests are in wireless communications, including MIMO, Cooperative communications, heterogeneous networks, coding and OFDM.

 <https://orcid.org/0000-0002-6484-0055>

E-mail: ibrahim_aljebori@yahoo.com

Electrical Engineering

Tikrit University

Salahaddin, Iraq



Ayad Atiyah Abdulkafi received his B.Sc. and M.Sc. degrees in Electrical Engineering (majoring in Telecommunications) from Al-Mustansiriya University, Baghdad, Iraq, in 2001 and 2004, respectively. He then received his Ph.D. in Wireless Communication Engineering from University Tenaga National

(UNITEN), Malaysia in 2015. He was a postdoctoral research Fellow at the Faculty of Engineering, Multi-

media University, Cyberjaya, Malaysia. He is a staff member at the Faculty of Electrical Engineering, Tikrit University, Iraq. His research interests are in wireless communications, including, LTE, heterogeneous networks, energy-efficient wireless network design, Green Cellular Networks, OFDM, optical wireless communications, and Visible Light Communication.

 <https://orcid.org/0000-0002-1160-6011>

E-mail: al.ayad@yahoo.com

Electrical Engineering

Tikrit University

Salahaddin, Iraq



Mohamed Khalf Hussein received his B.Sc. degree in Electronics Engineering from Air-Military Engineering Academy, Faculty of Electronics, Sarajevo, Yugoslavia, in 1982 and M.Sc. in Electronics Engineering, from the University of Belgrade, Electrical Engineering College, Belgrade, Yugoslavia, in 1984. He is currently a staff

member at the Faculty of Electrical Engineering, Shirqat, Tikrit University, Iraq. His research interests are in wireless communications, optical communications networks, coding, and OFDM.

E-mail: muh_kh48@yahoo.com

Electrical Engineering

Tikrit University

Salahaddin, Iraq



Saad Mshhain Hardan received his B.Sc. and M.Sc. degrees in Electrical and Electronic Engineering, Communication Engineering from the University of Technology, Baghdad, Iraq in 1997 and 2005, respectively. He is a staff member at the College of Engineering, Tikrit University, Iraq. He is currently a Ph.D.

candidate at Altinbas University, Turkey. His research interests are in Radio Wireless Communication, Optical Wireless Communication, Visible Light Communication, Digital Communication, and Wireless Communication Networks.

 <https://orcid.org/0000-0003-2395-2525>

E-mail: saadmhf@yahoo.com

Electrical Engineering

Tikrit University

Salahaddin, Iraq

A Modified Approach to Punctured Product Polar Codes

Mohammad Abu Hanif and Sina Vafi

College of Engineering, IT and Environment, Charles Darwin University, Darwin, Australia

<https://doi.org/10.26636/jtit.2019.132219>

Abstract—A high-performance polar code introduced as a product polar code can be constructed by concatenating two short length polar codes. The punctured structure of this code was achieved by puncturing one of its constituent codes. The constructed polar code provided better performance than a single polar code in the error floor region. However, its performance in the waterfall region was not as good as that of single polar codes. This paper proposes a new puncturing algorithm for product codes constructed by two identical polar codes. Puncturing is conducted on both constituent codes, to ensure that the new code outperforms the previously punctured product polar and single polar codes. This is evident in both waterfall and error floor regions.

Keywords—polar codes, product codes, punctured codes, iterative decoding.

1. Introduction

Polar codes are recently invented promising forward error correction (FEC) codes [1], which are well known for their low complexity in encoding and decoding. However, the lengths of codes are not flexible as they are in order of 2^n ($n = 1, 2, \dots$). This limits their application in video services and control-based systems, which usually require unequal error protection in their structure. Similar to other linear block codes, obtaining an arbitrary length for polar codes is possible by puncturing a number of bits from its original codeword.

Several puncturing methods were proposed, which are mainly applicable for single polar codes (SPC). In one of the methods, an exhaustive search is performed to delete rows and columns that maximize the exponent of the reduced generator matrix $\mathbf{G}$ [2]. This technique offers better performance compared to random puncturing. However, implementation of an exhaustive search algorithm for determining a good polarized matrix still remains a time-consuming process. Another class of punctured polar codes was proposed in [3], where the first η_p unreliable bits in the bit-reversal order are punctured. An improvement is presented by Wang *et al.* [4], where a puncturing pattern is obtained based on the column weights of the generation matrix. In [5], puncturing and shortening sets are formed based on bit-reversal permutation of the most unreliable bits. Similarly, a technique was proposed to ensure short polar codes, which do not select consecutive coded bits

as punctured ones [6]. Finally, limitations of the above-mentioned techniques are overcome by puncturing bits at different levels of the code [7].

In general, performance of SPCs is limited due to the existence of a low minimum weight, or a relatively high minimum weight with high multiplicity. As a solution, product polar codes (PPCs) were proposed, which are constructed by the serial combination of two half-rate polar codes [8]. These codes generate the same minimum weight but with a lower multiplicity compared to SPCs with rate 1. It is also shown that with the same code rates and lengths, punctured PPCs provide better performance than punctured SPCs. The improvement is mainly evident in the medium to high signal-to-noise ratios. In the proposed method, puncturing of the product code is only conducted at one of its constituent codes. Low performance of PPCs in the waterfall region is mainly related to the weak performance of their punctured constituent codes [8].

This paper presents an algorithm, which improves the performance of punctured product polar codes, where their two constituent codes are punctured. In this way, bits having the highest error probability are punctured from the constituent codes. This ensures that the most unreliable bits are punctured for the product polar code and allows its constituent decoders to efficiently recover the original information.

The rest of the paper is organized as follows. A brief introduction of polar codes is given in Section 2. In Section 3, structure of product polar codes and a new puncturing technique applied for both constituent codes are presented. The effects that different puncturing techniques exert on iterative decoding are analyzed in Section 4. Numerical results are given in Section 5. Finally, Section 6 concludes the paper.

2. Polar Codes

The generator matrix, of a (N, K, d_{min}) polar code with length N , dimension K and minimum weight (d_{min}) is formed by taking n -th, $n = \log_2(N)$ Kronecker power of a binary matrix $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, which is given by $G_N = B^{\otimes n}$.

Codewords of these codes are formed by multiplying a message vector of length K with a sub-matrix of G_N , which includes only K rows of the generator matrix (G_K).

These rows are selected based on the reliability of N -th bit-channel. For any binary discrete memoryless channel, the reliability of N -th bit-channels can be recursively constructed by [1]:

$$Z(w_{2i}^j) = \begin{cases} 2Z(w_i^j) - Z(w_i^j)^2 & \text{if } 1 \leq j \leq i \\ Z(w_i^{j-i})^2 & \text{if } i+1 \leq j \leq 2i \end{cases}, \quad (1)$$

where i is a positive integer and in order of 2 ($1 \leq i \leq 2^{n-1}$) and $Z(w_{2i}^j)$ denotes the Bhattacharyya parameter of j -th bit-channel at stage $2i$ ($2i \leq N$). The code construction method shown in Eq. (1) was initially designed for the binary erasure channel (BEC), when the equation is initialized by $Z(w_1^1) = \frac{1}{2}$. For the binary additive white Gaussian noise (AWGN) channel, with zero mean and variance $\frac{N_0}{2}$, $Z(w_1^1)$ can be given by [9], [10]:

$$Z(w_1^1) = e^{-\frac{E_c}{N_0}}, \quad (2)$$

where $E_c = RE_b$, R is the rate of the code and E_b is the energy spent per bits.

The smallest $Z(w)$ value obtained from Eq. (1) refers to the most reliable bit-channel, known as a good bit-channel. Let A_I represents a set of good bit-channels. The remaining bits are known as bad bit-channels and are denoted as elements of A_I^c . Good channels are suitable for carrying information, and the rest are recognized as frozen bits.

Let $\mathbf{u} = (u_1 \ u_2 \ \dots \ u_K)$ be a message vector of length K and $\mathbf{x} = (x_1 \ x_2 \ \dots \ x_N)$ be the corresponding coded sequence, which is generated by $\mathbf{x} = \mathbf{u}\mathbf{G}_K$, $K \in \{A_I\}$. This paper follows the systematic encoding and decoding proposed by Arikan in [13]. The d_{min} of a systematic polar code is determined based on weights of its rows, whose indices are elements of A_I . This is given by [14], [15]:

$$d_{min}(C^{SPC}) = \min_{b \in A_I} \{2^{\omega(b)}\}, \quad (3)$$

where $\omega(b)$ is the number of ones in the binary expansion of b .

Based on an exhaustive search algorithm, it is proven that some polar codes have a low minimum weight compared to their length. On the other hand, some codes have high multiplicity for their relatively high minimum weights. These specifications deteriorate performance of codes at

medium to high signal-to-noise ratios, which is mainly dominated by their minimum weight specifications. By contrast, with the same code rate and length, it is possible to construct product codes whose minimum weights have lower multiplicity than those obtained from single polar codes [8].

3. Product Polar Codes (PPCs)

In (N, K) product polar codes with rate R , every message with length $K = K_1 \times K_2$ is converted to K_2 blocks with length K_1 and encoded by the half rate ($R_1 = \frac{K_1}{N_1} = \frac{1}{2}$) outer (N_1, K_1, d_{min_1}) polar code. These codewords $(\mathbf{x}_{K_2 \times N_1}^1)$ are then encoded by another half rate inner (N_2, K_2, d_{min_2}) polar code to form a two-dimensional codeword matrix $(\mathbf{x}_{N_1 \times N_2})$ of (N, K, d_{min}) product polar codes with $R = \frac{1}{4}$, where $N = N_1 \times N_2$ and $d_{min} = d_{min_1} \times d_{min_2}$.

Interactive iterative decoding is used to decode the information received from the channel. The basic structure of iterative decoding of PPCs is shown in Fig. 1. Here, $\mathbf{y} = [\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \ \mathbf{y}_{N_1}]^T$ denotes the noisy received sequence, where $\mathbf{y}_l = [y_{l,1} \ y_{l,2} \ \dots \ y_{l,N_2}]$ is the received block at time l . In the figure, $L_{em}(\hat{\mathbf{x}}_l)$ s is the extrinsic information of the estimated codeword $(\hat{\mathbf{x}}_l)$ delivered by the m -th decoder ($m = 1, 2$). At t -th iteration, log-likelihood ratios (LLRs) of the decoders are determined by:

$$L_1^{(t)}(\hat{\mathbf{x}}_\alpha) = Le_2^{\pi(t-1)}(\hat{\mathbf{x}}_\alpha) + \mathbf{y}_\alpha \quad (4)$$

$$L_1^{(t)}(\hat{\mathbf{x}}_\beta) = Le_1^{(t-1)}(\hat{\mathbf{x}}_\beta)$$

$$L_2^{(t)}(\hat{\mathbf{x}}_\alpha) = Le_1^{\pi^{-1}(t)}(\hat{\mathbf{x}}_\alpha) + \mathbf{y}_\alpha^{\pi^{-1}},$$

where $\alpha \in \{A_I\}$, $\beta \in \{A_I^c\}$ and $Le_2^{(0)}(\hat{\mathbf{x}}_\alpha) = \mathbf{0}$. Here, $(o)^\pi$ and $(o)^{\pi^{-1}}$ are the interleaved and de-interleaved information of (o) . Note that a conventional concatenation technique is followed in this paper, where the inner code cannot be smaller than the outer code. This means $N_2 \geq N_1$ and $K_2 \geq K_1$.

Although product polar codes reduce the multiplicity of d_{min} , they are represented as low rate codes. In order to construct product codes with a rate greater than $\frac{1}{4}$, a proper puncturing technique is required to maintain a high level of error correcting capability.

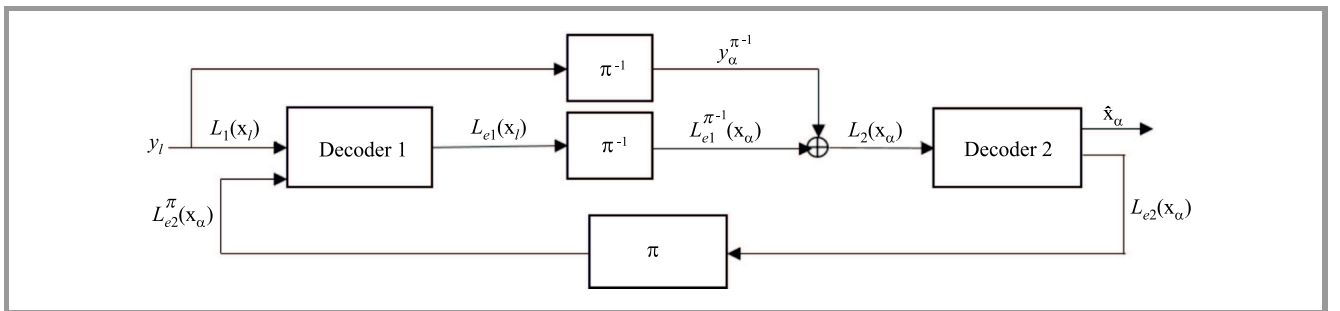


Fig. 1. Decoder for a two-dimensional product polar code.

3.1. Puncturing of Product Polar Codes

The arbitrary lengths of product polar codes can be obtained by puncturing a number of bits from codewords. Puncturing can only be performed at the inner code. In this case, codewords of the punctured product polar code are constructed by deleting n_ϕ rows and columns from the generator matrix of $(N_2, K_2, d_{\min_2})$ code [8]. This concludes a $(N_1(N_2 - n_\phi), K_1.K_2)$ code, whose minimum weight ($d'_{\min}$) is upper bounded with the minimum weight of $d_{\min_1} \cdot (d_{\min_2} - (n_1)_{n_\phi})$, where $(n_1)_{n_\phi}$ gives the number of punctured bits with the value of 1. Obviously, punctured bits are the most unreliable bits of $(N_2, K_2, d_{\min_2})$ polar code, which are also elements of A_I^c .

Example 1: A (64, 16) product polar code can be formed by concatenating two (8, 4, 4) polar codes, $N_1 = N_2 = 8$ and $K_1 = K_2 = 4$.

Let

$$\mathbf{u} = \begin{pmatrix} u_{1,1} & u_{1,2} & u_{1,3} & u_{1,4} \\ u_{2,1} & u_{2,2} & u_{2,3} & u_{2,4} \\ u_{3,1} & u_{3,2} & u_{3,3} & u_{3,4} \\ u_{4,1} & u_{4,2} & u_{4,3} & u_{4,4} \end{pmatrix}$$

be the message with length K_1 , which is encoded by the outer (8, 4, 4) polar code. Its two-dimensional codeword matrix is given by:

$$\mathbf{x}^1 = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} & u_{1,1} & p_{1,4} & u_{1,2} & u_{1,3} & u_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & u_{2,1} & p_{2,4} & u_{2,2} & u_{2,3} & u_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,3} & u_{3,1} & p_{3,4} & u_{3,2} & u_{3,3} & u_{3,4} \\ p_{4,1} & p_{4,2} & p_{4,3} & u_{4,1} & p_{4,4} & u_{4,2} & u_{4,3} & u_{4,4} \end{pmatrix}, \quad (5)$$

where $p_{i,j}$ are the parity bits and $i, j \in \{A_I^c\}$. The $\mathbf{x}^1$ is interleaved and encoded by the inner code. The inner code is punctured by $\eta_\phi = 2$ bits. As the first and fifth bit-channels, according to Eq. (1), are the most unreliable bits amongst the eight bit-channels of (8, 4, 4) the polar code, every message of $\mathbf{x}^1$ is encoded by removing first and fifth rows and columns of $\mathbf{G}_8$ matrix. This concludes a two-dimensionally punctured codeword presented in below:

$$\mathbf{x} = \begin{pmatrix} \times & \times & \times & \times & \times & \times & \times & \times \\ p'_{2,1} & p'_{2,2} & p'_{2,3} & p'_{2,4} & p'_{2,5} & p'_{2,6} & p'_{2,7} & p'_{2,8} \\ p'_{3,1} & p'_{3,2} & p'_{3,3} & p'_{3,4} & p'_{3,5} & p'_{3,6} & p'_{3,7} & p'_{3,8} \\ p_{1,1} & p_{1,2} & p_{1,3} & u_{1,1} & p_{1,4} & u_{1,2} & u_{1,3} & u_{1,4} \\ \times & \times & \times & \times & \times & \times & \times & \times \\ p_{2,1} & p_{2,2} & p_{2,3} & u_{2,1} & p_{2,4} & u_{2,2} & u_{2,3} & u_{2,4} \\ p_{3,1} & p_{3,2} & p_{3,3} & u_{3,1} & p_{3,4} & u_{3,2} & u_{3,3} & u_{3,4} \\ p_{4,1} & p_{4,2} & p_{4,3} & u_{4,1} & p_{4,4} & u_{4,2} & u_{4,3} & u_{4,4} \end{pmatrix}, \quad (6)$$

where $p'_{i,j}$, $i, j \in \{A_I^c\}$ are the parity bits produced by the inner code. Indeed, all 8 codewords expressing columns of the matrix are punctured by two bits. This sums up

a (48, 16) product polar code formed by combining (8, 4, 4) and (6, 4, 4) polar codes.

3.2. The Modified Approach

A new puncturing scheme for the product polar code is constructed by puncturing both constituent codes. Indeed, a $(N', K_1.K_2)$ punctured code is formed, where $N' = (N_1 - n'_{\phi_1})(N_2 - n'_{\phi_2})$. In this case, n'_{ϕ_1} and n'_{ϕ_2} are the numbers of unreliable bits punctured from $(N_1, K_1, d_{\min_1})$ and $(N_2, K_2, d_{\min_2})$ polar codes, respectively.

Figure 2 shows steps in the process of construction of (49, 16) product polar code formed by two punctured (8, 4, 4) polar codes. The bit-channel one is the most unreliable amongst the eight channels of (8, 4, 4) polar code. Therefore, encoding at the outer code is performed by removing the first row and column of its generator matrix. Then, every message of $\mathbf{u}$ (as shown in Fig. 2) is encoded to form a punctured codeword of the outer code. As bits in the first column are already punctured, there are $l-1$ ($l = N_1$) bits remaining in each codeword.

Let $\mathbf{x}^1$ be the matrix formed by punctured codewords of the outer code. This matrix is interleaved and each of its columns is used as the message for the inner code. Hence, a two-dimensional codeword matrix ($\mathbf{x}$) at the output of inner encoder is obtained. One may observe that only one bit from $l-1 = 7$ codewords of inner (8, 4, 4) code is punctured. Again, for this code, the first bit of the codeword is the most unreliable one. Therefore, bits positioned at the first row of $\mathbf{x}$ are punctured to conclude a (49, 16) code.

The minimum weights ($d''_{\min}$) of these punctured codes are given by $(d_{\min_1} - (n'_1)_{n'_{\phi_1}}) \cdot (d_{\min_2} - (n'_2)_{n'_{\phi_2}})$, where $(n'_1)_{n'_{\phi_1}}$ and $(n'_2)_{n'_{\phi_2}}$ are numbers of bits (with the value of 1) punctured from $(N_1, K_1, d_{\min_1})$ and $(N_2, K_2, d_{\min_2})$ polar codes, respectively. If $(n)_{n_\phi} = (n'_1)_{n'_{\phi_1}} + (n'_2)_{n'_{\phi_2}}$ and $(n'_1)_{n'_{\phi_1}} = (n'_2)_{n'_{\phi_2}}$, $d''_{\min} = d'_{\min} + (n'_1)_{n'_{\phi_1}} \cdot (n'_2)_{n'_{\phi_2}}$. Obviously, $d''_{\min} > d'_{\min}$. This shows that the product code punctured on the basis of two constituent codes has better performance than codes whose puncturing is only conducted on one of their constituent codes. Based on this analysis, improvement is primarily expected at the medium to high signal-to-noise ratios, where the codes' performance is dominated by their minimum weight specifications.

4. Puncturing Effect on the Iterative Decoding of PPCs

The effect of puncturing is verified through the iterative decoding process. In each iteration, LLRs of the information are updated and fed back to the input of the decoder until a maximum number of iterations ($\text{ite}_{\max}$) is reached. At t -th iteration ($t \leq \text{ite}_{\max}$), the iterative decoding process can be terminated if the following assumptions on the extrinsic information at the output of two decoders are satisfied [11]:

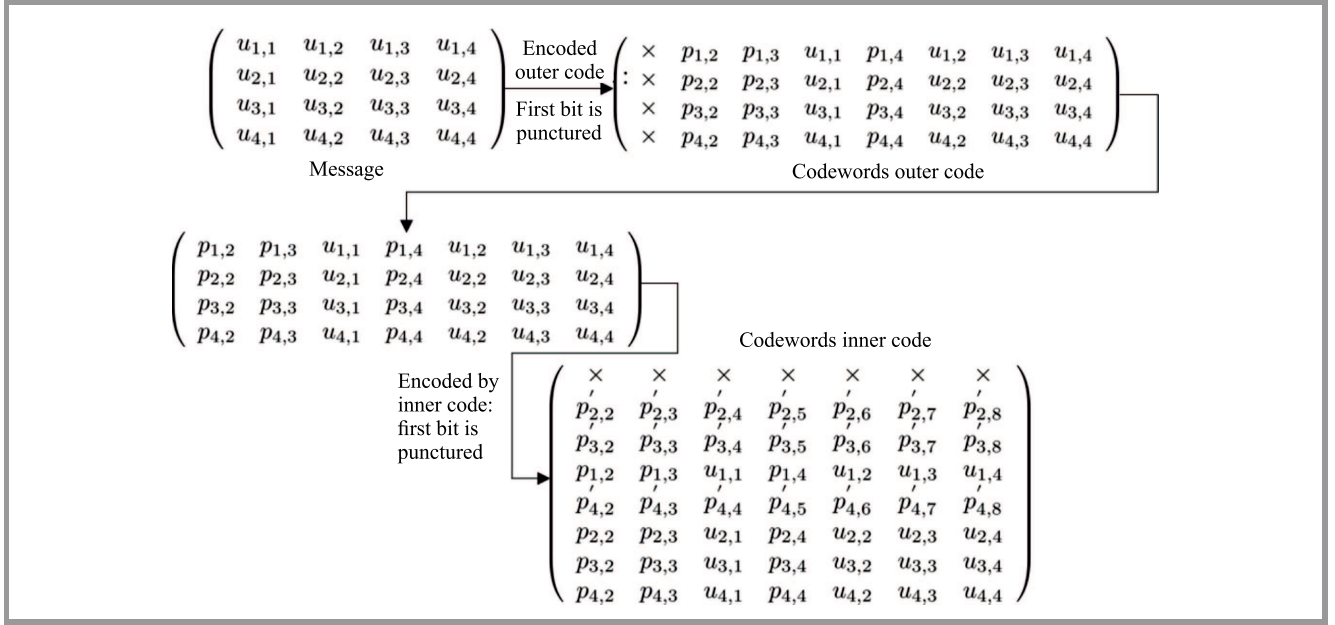


Fig. 2. Puncturing of two (8, 4, 4) polar codes for constructing a (49, 16) PPC.

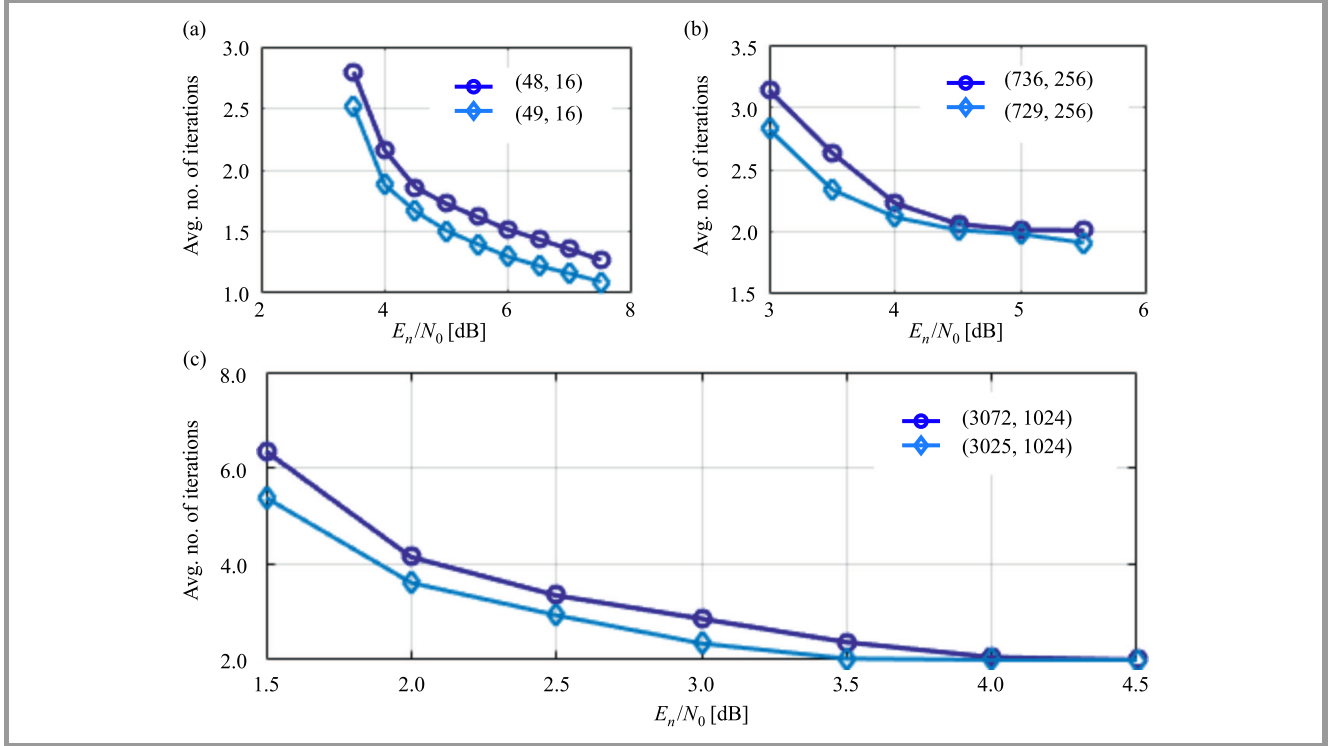


Fig. 3. Average number of iterations for punctured product polar codes.

- The polarities of soft-output information obtained from both decoders are not changed at two consecutive iterations. This is done by checking the sign of LLRs of two decoders, which is given by:

$$\begin{aligned} \text{sign}(L_{e1}^{(t)}(\hat{x}_{i,j})) &= \text{sign}(L_{e2}^{(t)}(\hat{x}_{i,j})) \\ &= \hat{x}_{i,j}^{(t)} = [\pm 1], \end{aligned} \quad (7)$$

where $\hat{x}_{i,j}$ denotes the estimate value of $x_{i,j}$ and $i, j \in \{A_I\}$.

- The difference between the magnitudes of the extrinsic information of a decoder input and output is small.

Figure 3 shows number of iterations applied for iterative decoding of different polar codes. In simulations, AWGN

channel is modeled and the code is modulated by binary phase shift keying (BPSK). The pipeline decoder with the maximum of 60 iterations is constructed for the decoding of constituent polar codes. Five iterations are also applied in process of interactive iterative decoding between the two decoders.

For (48, 16) and (49, 16) PPCs, puncturing is conducted on one and two of their constituent codes, respectively. In this case, two bits from each of eight codewords are punctured for (48, 16) code, while only one bit of seven codewords from each of (8, 4, 4) codes is punctured for the (49, 16) code. As the (49, 16) product code applied a lower number of punctured bits in its codewords, a weaker impact of the punctured bits on the LLR calculation of the information bits is observed. While the most unreliable bits are punctured from the inner code of (48, 16) PPC, some bits recognized as reliable bits for the outer code also punctured. Due to the removal of these reliable bits, interactive iterative decoding does not provide a good error correcting capability. Hence, decoding of this code is conducted by a higher number of iterations than in the case of the (49, 16) code.

Similarly, (736, 256) and (729, 256) PPCs are formed by puncturing one and both of their constituent codes, respectively. In this case, $N_1 = N_2 = 32$, $n_\phi = 9$ and $n'_{\phi 1} = n'_{\phi 2} = 5$. This means that nine bits are punctured from each of 32 codewords, when one constituent code is only punctured. Alternatively, five bits are punctured from 27 codewords, when both of the constituent codes are punctured. Although a higher number of bits are punctured to obtain the (729, 256) code, its iterative decoder applies a lower number of iterations than that of the (736, 256) code. Moreover, 16 bits are punctured from each of 64 codewords of the (3072, 1024) code, while nine bits are punctured from 55 codewords of the (3025, 1024) code. As shown in the figure, a lower number of iterations for the (3025, 1024) code is applied compared to (3072, 1024) at $\frac{E_b}{N_0} \leq 3.5$ dB.

5. Numerical Results

In this section, the block error rate (BLER) and bit error rate (BER) performance of product codes under AWGN channel with BPSK modulation are evaluated. Codes are decoded by using the belief propagation decoding method. Again, for product codes, the iterative decoding of their constituent decoders is conducted with 60 iterations. Five iterations are considered for the interactive decoding between two constituent decoders. For SPCs, decoding is also accomplished by 60 iterations.

Figure 4 shows performance of (48, 16) and (49, 16) product codes, as well as the (49, 16) single polar code. As mentioned in previous sections, (48, 16) and (49, 16) product codes are formed by puncturing one and both of their constituent codes, respectively. It is observed that (49, 16) product code has a similar performance to that of the single punctured code, while it performs better than (48, 16) product code. In addition, at $\frac{E_b}{N_0} \geq 5.5$ dB,

(49, 16) product code shows a performance better than that of (49, 16) SPC, as its error floor occurs at $\text{BER} < 10^{-5}$. The confidence interval of (49, 16) product code at $\frac{E_b}{N_0} = 6.0$ dB is $[0.1276 \times 10^{-3}, 0.3915 \times 10^{-3}]$, which is $[0.1907 \times 10^{-3}, 0.4945 \times 10^{-3}]$ for the SPC.

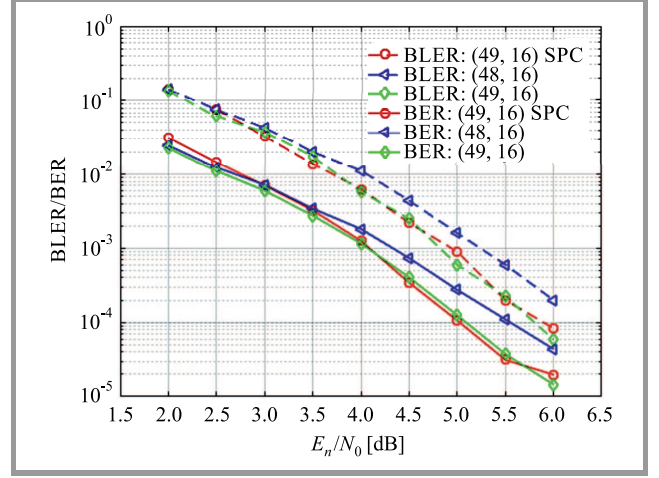


Fig. 4. BLER and BER performance comparison of (48, 16) and (49, 16) product polar codes and (49, 16) SPC.

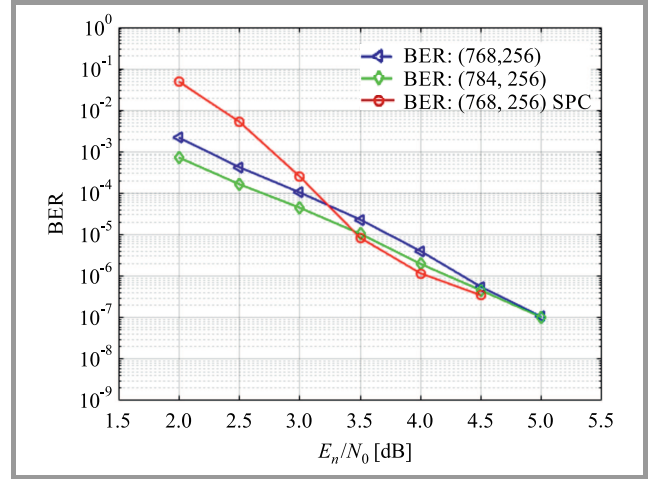


Fig. 5. BER performance comparison of (768, 256) and (784, 256) product polar codes and (768, 256) SPC.

Figure 5 shows the performance of (768, 256) and (784, 256) product codes, as well as of the (768, 256) single polar code. The (768, 256) product code is constructed by con-catenating (32, 16, 4) and (24, 16, 4) polar codes, whereas (784, 256) product code is formed by combining two (28, 16, 4) polar codes. In the (768, 256) product code, $\eta_\phi = 8$ bits from every codeword are punctured, whereas half of η_ϕ bits are punctured from 28 codewords of the (784, 256) product code. The effect of interactive decoding is observed at $\frac{E_b}{N_0} \geq 4.5$ dB for these codes, when the decoder can properly recover the punctured information from a block of information. As a lower number of bits is punctured from codewords of the (784, 256) product code, this code provides a better performance than the (768, 256)

product code for $2.0 \text{ dB} \leq \frac{E_b}{N_0} \leq 4.0 \text{ dB}$. The (784, 256) product code also provides better BER performance than (768, 256) single polar code for $\frac{E_b}{N_0} \geq 4.5 \text{ dB}$. Figure 3b shows also that a lower number of iterations is required for this product code. By contrast, for the single polar code, 256 bits are removed from the codeword, which means that its error floor is at $\frac{E_b}{N_0} \geq 4.5 \text{ dB}$. No error floor has been observed for the product code.

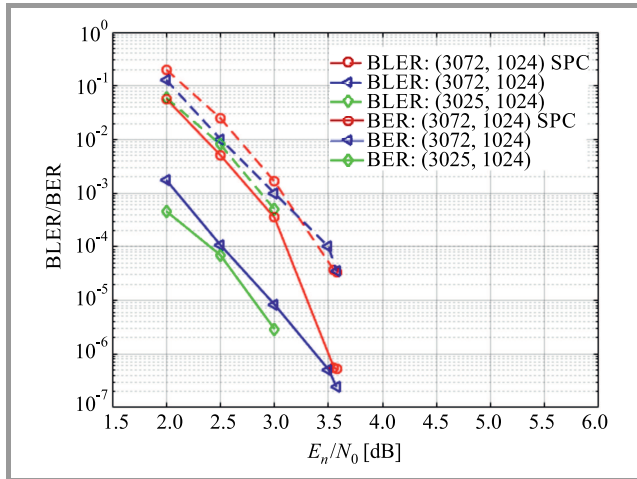


Fig. 6. BLER and BER performance comparison of (3072, 1024) and (3025, 1024) product polar codes, and (3072, 1024) SPC.

Figure 6 shows the BER and BLER performance of (3072, 1024) and (3025, 1024) codes. It is seen that both product codes have a better performance compared to that of the single code at the waterfall region. This is achieved for $\frac{E_b}{N_0} \leq 3.0 \text{ dB}$. For $\frac{E_b}{N_0} > 3.0 \text{ dB}$, (3072, 1024) SPC and PPC formed by puncturing of one of their constituent codes demonstrate a similar performance. At the same time, their confidence intervals are $[0.4404 \times 10^{-3}, 0.7497 \times 10^{-3}]$ and $[0.3970 \times 10^{-3}, 0.6932 \times 10^{-3}]$, respectively. However, the (3025, 1024) PPC constructed based on the puncturing of two constituent codes outperforms them.

6. Conclusions

The paper proposed a framework of product punctured polar codes, where puncturing is applied to both of the constituent codes. The analysis and simulations concluded that codes constructed by this method outperform the previously proposed ones, with a lower number of iterations applied in their iterative decoding structure.

References

- [1] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels", *IEEE Trans. on Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009 (doi: 10.1109/TIT.2009.2021379).
- [2] D. M. Shin, S. C. Lim, and K. Yang, "Design of length-compatible polar codes based on the reduction of polarizing matrices", *IEEE Trans. on Commun.*, vol. 61, no. 7, pp. 2593–2599, 2013 (doi: 10.1109/TCOMM.2013.052013.120543).

- [3] K. Niu, K. Chen, and J.-R. Lin, "Beyond turbo codes: Rate-compatible punctured polar codes", in *Proc. IEEE Int. Conf. on Commun. ICC 2013*, Budapest, Hungary, 2013, pp. 3423–3427 (doi: 10.1109/ICC.2013.6655078).
- [4] R. Wang and R. Liu, "A novel puncturing scheme for polar codes", *IEEE Commun. Lett.*, vol. 18, no. 12, pp. 2081–2084, 2014 (doi: 10.1109/LCOMM.2014.2364845).
- [5] I. L. V. Bioglio and F. Gabry, "Low-complexity puncturing and shortening of polar codes" [Online]. Available: <https://arxiv.org/abs/1701.06458v1>.
- [6] X. W. Q. Y. L. Zhang, Z. Zhang, and Y. Chen, "On the puncturing patterns for punctured polar codes", in *Proc. IEEE Int. Symp. on Inform. Theory ISIT 2014*, Honolulu, HI, USA, 2014, pp. 121–125 (doi: 10.1109/ISIT.2014.6874807).
- [7] M. Hanif and S. Vafi, "An efficient puncturing method for the short and long length polar codes", in *Proc. 11th Int. Conf. on Sig. Process. and Commun. Syst. ICSPCS 2017*, Surfers Paradise, QLD, Australia, 2017 (doi: 10.1109/ICSPCS.2017.8270452).
- [8] M. A. Hanif and S. Vafi, "A novel scheme of product polar codes with high efficiency and rate compatibility" (under review at *Wireless Communications and Mobile Computing*).
- [9] H. Li and J. Yuan, "A practical construction method for polar codes in AWGN channels", in *Proc. IEEE TENCON Spring Conf.*, Sydney, NSW, Australia, pp. 223–226, 2013 (doi: 10.1109/TENCONSpring.2013.6584444).
- [10] H. Vangala, E. Viterbo, and Y. Hong, "A comparative study of polar code constructions for the AWGN channel" [Online]. Available: <https://arxiv.org/abs/1501.02473>
- [11] R. Y. Shao, S. Lin, and M. P. C. Fossorier, "Two simple stopping criteria for turbo decoding", *IEEE Trans. on Commun.*, vol. 47, no. 8, pp. 1117–1120, 1999 (doi: 10.1109/26.780444).
- [12] E. Arikan, "Polar codes: A pipelined implementation", in *Proc. 4th Int. Symp. on Broadcast Commun. ISBC 2010*, Melaka, Malaysia, 2010, pp. 11–14 [Online]. Available: <https://pdfs.semanticscholar.org/a9ed/49203bffe7b124984a5874fba0940103d8ab.pdf>
- [13] E. Arikan, "Systematic polar coding", *IEEE Commun. Lett.*, vol. 15, no. 8, pp. 860–862, 2011 (doi: 10.1109/LCOMM.2011.061611.110862).
- [14] N. Hussami, S. B. Korada and R. Urbanke, "Performance of polar codes for channel and source coding", in *Proc. IEEE Int. Symp. on Inform. Theory ISIT 2009*, Seoul, South Korea, 2009 (doi: 10.1109/ISIT.2009.5205860).
- [15] Q. Zhang, A. Liu, Y. Zhang, and X. Liang, "Practical design and decoding of parallel concatenated structure for systematic polar codes", *IEEE Trans. on Commun.*, vol. 64, no. 2, pp. 456–466, 2016 (doi: 10.1109/TCOMM.2015.2502246).



Mohammad Abu Hanif is currently pursuing his Ph.D. degree in Engineering and Information Technology at Charles Darwin University, Australia. He holds a M.Sc. degree in Information and Communications from South Korea and a B.Sc. degree in Computer Science and Engineering from Bangladesh. After his master course, he worked on

several projects as a Research Assistant in South Korea. His research interests are in information theory, error-control coding and wireless communications.

E-mail: mohammad.hanif@cdu.edu.au

College of Engineering, IT and Environment
Charles Darwin University
Darwin, Australia



Sina Vafi received his Ph.D. degree from the University of Wollongong (UoW), Wollongong, Australia, in 2006. In 2007, he joined Charles Darwin University as a Lecturer of Electrical and Electronics Engineering. In 2008, he became a Senior Lecturer. From 2006 to 2007, he was involved in re-

search, design, and implementation of 2.4 GHz broadband wireless mesh networks at UoW. He currently leads research projects on error protection techniques and their application for broadband wireless video transmission systems. His main research interests include algebraic coding theory and video transmission systems.

E-mail: sina.vafi@cdu.edu.au

College of Engineering, IT and Environment
Charles Darwin University
Darwin, Australia

Ku Band Bethe Hole Coupler Using Gap Waveguide Technology

Efat Nematpour, Mohamad Hossein Ostovarzadeh, and Seyed Ali Razavi

Graduate University of Advanced Technology, Kerman, Iran

<https://doi.org/10.26636/jtit.2019.132119>

Abstract—The gap waveguide technology is a new technique used for designing and fabricating microwave components, ensuring a low-loss and easy fabrication process, especially at high frequencies, and allowing for the production of multi-layer structures due to the lack of requirement of an electrical connection between the metal layers of the waveguide structure. This paper presents the design and areas of implementation of single-hole and multi-hole 20 dB Bethe couplers, using the groove gap waveguide (GGW) technology for Ku band. Simulation results show that the operating bandwidth of the proposed design is over 40% wider, and its isolation rate is more than 25 dB higher. By using the multi hole configuration, a bandwidth that is more than 59% wider and the isolation rate of over 30 dB may be obtained.

Keywords—Bethe hole coupler, multi-hole coupler, groove gap waveguide.

1. Introduction

Microwave components are currently designed using conventional metallic waveguides, microstrip lines, strip lines and SIW. Such technologies suffer from certain disadvantages that limit their application. Conventional waveguides offer the lowest loss rates and the highest power handling capabilities. However, the process of their fabrication is expensive challenging at high frequencies, due to the need to provide a fine electrical contact between metal plates. Microstrip and coplanar lines are a low-cost solution that is suitable for integrating active microwave components on a PCB, but they suffer from a high insertion loss rate at high-frequencies, and offer a low power handling capability. SIW structures are compact, planar and generate no undesired radiation, but they suffer from high dielectric and radiation losses at millimeter wave frequency ranges, due to the discrete nature of sidewalls which do not provide a perfect shielding [1]–[3].

The gap waveguide technology has recently gained in popularity due to its easy manufacturing combined with the lack of the requirement to ensure an electric contact between the metal plates, especially at millimeter-wave frequencies. The other advantages of the solution include the following: easy assembly (especially for multi-layer configurations) and low loss rates due to the lack of presence of a dielectric material in the propagation path, as well as no parasitic radiation and high-quality rates. So far,

several microwave components based on the gap waveguide technology have been implemented, including various types of antennas [4]–[9], power combiners [10], diplexer-power dividers [11] and filters [12], [13]. Four different types of gap waveguide structures include the following: ridge gap waveguide (RGW) [13], groove gap waveguide (GGW) [14], inverted microstrip gap waveguide [15] and microstrip-ridge gap waveguide [16].

The purpose of this work is to design a Bethe hole coupler using the groove gap waveguide technology. In Section 2, we discuss the GGW structure. The design and simulation results of single-hole and multi-hole Bethe couplers are presented in Sections 3 and 4, respectively. In Section 5, the performance of both couplers is compared. Finally, in Section 6, concluding remarks are presented.

2. Groove Gap Waveguide Structure

GGW technology is based on the propagation of waves in the gap between two parallel metallic surfaces providing a non-metal contact-guiding structure [14]. Depending on the geometry of the groove, it is possible to use vertical (GGW-VP) or horizontal (GGW-HP) polarizations [17],

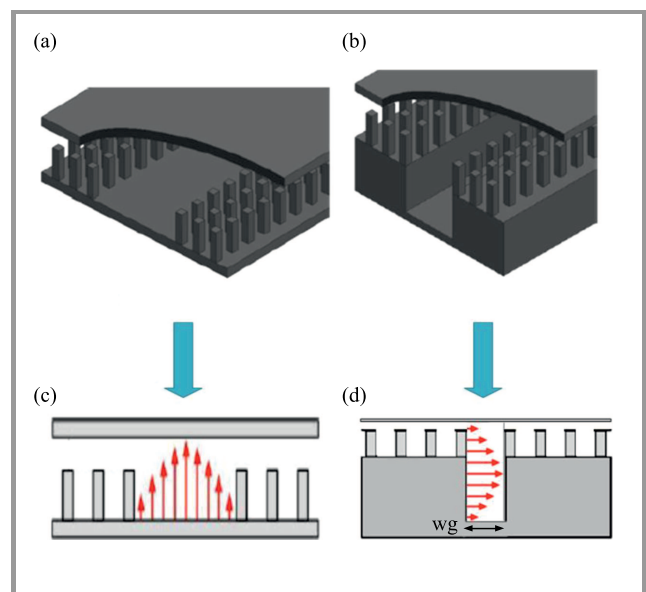


Fig. 1. GGW geometry: (a) GGW-VP, (b) GGW-HP, (c) E field distribution in GGW-VP, (d) E field distribution in GGW-HP.

as shown in Fig. 1. GGW is equivalent to a conventional rectangular waveguide that supports TE modes. There is no need to ensure electric contact between the top and bottom plates of the waveguide, which makes it suitable for designing microwave circuits operating at high frequencies. In this paper the GGW-VP version of the design is examined, based on the concept and design guidelines originating from [14].

3. Single-hole Bethe Coupler Design

The geometry of the proposed GGW Bethe hole coupler is shown in Fig. 2. It includes two layers of GGW which are coupled together through one circular hole located in the broad wall. 3D and 2D views of the structure are depicted in Fig. 2a and in Fig. 2b-c, respectively. The initial r_0 value for the hole diameter and its position relative to the center line of GGW are calculated using the formulas from [18] for the coupling coefficient of -20 dB. According to [19], zero thickness is considered for the coupling slot. Then, fine tuning is performed by full wave simulation of HFSS software to obtain optimum performance in terms of bandwidth, for the case in which the coupling slot has non-zero thickness.

The impact of slot thickness on the coupling coefficient is investigated for three frequencies within the operating

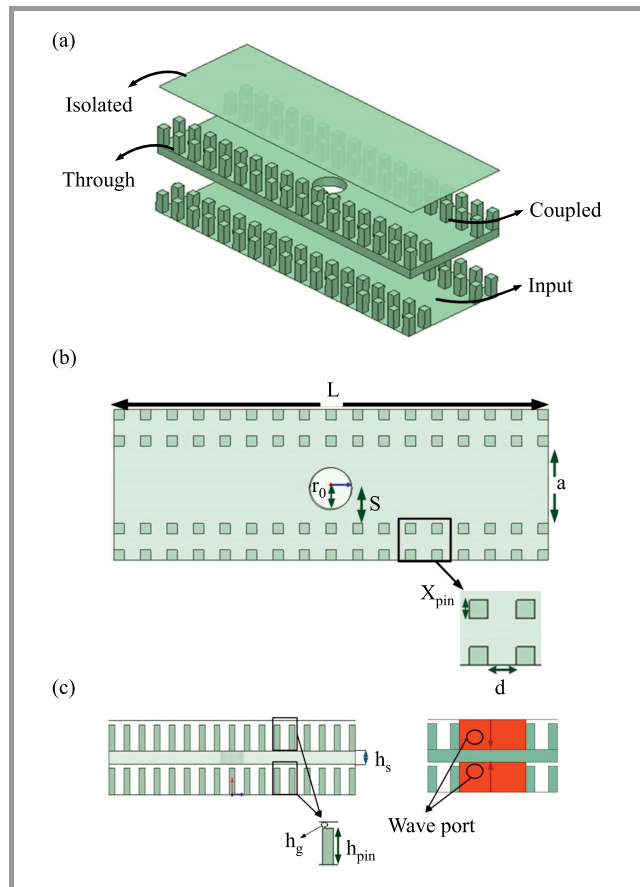


Fig. 2. GGW single-slot Bethe hole coupler structure: (a) the expanded 3D view, (b) top view, (c) side views.

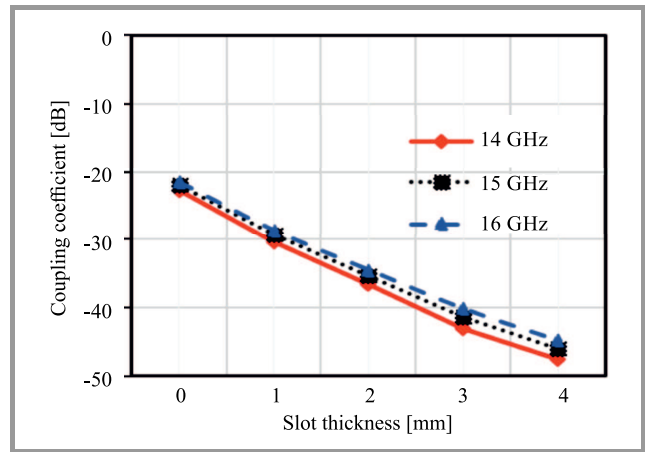


Fig. 3. The impact of slot thickness on the coupling coefficient ($r_0 = 2.7$ mm).

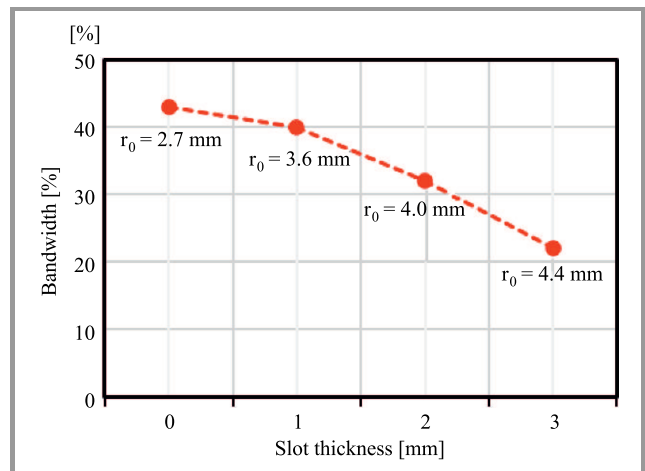


Fig. 4. The impact of slot thickness on bandwidth for the fixed coupling coefficient of -20 dB.

Table 1
Parameters of the designed coupler

Parameter	Value [mm]
a	15
d	3.1
h_g	1
h_{pin}	6
h_s	1
L	83.6
r_0	3.6
S	6.8
x_{pin}	2

bandwidth, as shown in Fig. 3. It can be observed that the coupling coefficient decreases as slot thickness increases. To compensate for this effect, the slot radius can be increased in order to enhance the coupling coefficient. The maximum bandwidth for the coupling coefficient -20 dB could be obtained using different slot thicknesses and cor-

responding slot radiiuses, as shown in Fig. 4. For a fixed coupling coefficient (-20 dB), as the slot thickness increases, a larger coupling slot is needed. However, the operating bandwidth decreases due to the increasing difficulty in matching impedance between upper and lower layers. Based on the results shown in Figs. 3 and 4, the coupling slot thickness of 1 mm was selected to ensure maximum bandwidth. The calculated dimensions of the designed 20 dB coupler are listed in Table 1. The coupler design was then simulated using HFSS software, and the results are compared with those obtained means of CST software.

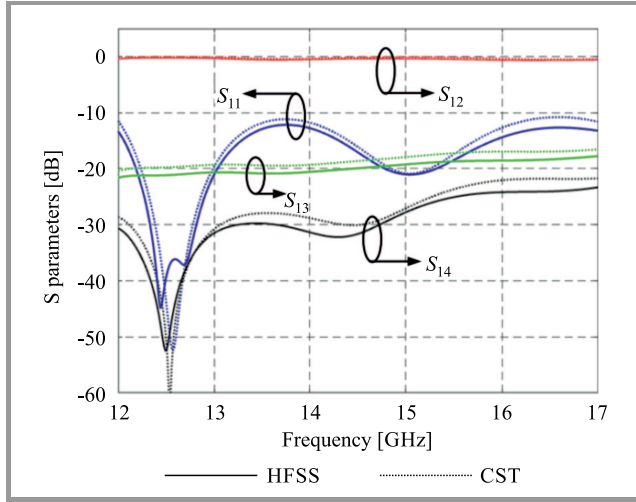


Fig. 5. S parameters of the designed coupler.

The simulation results of the designed single-hole Bethe coupler are shown in Fig. 5. S_{11} is below -10 dB over the entire bandwidth (12–17 GHz), while S_{12} is above -0.5 dB. The coupling coefficient C is about -20 ± 2 dB, and the isolation rate is better than -25 dB over the entire bandwidth. Table 2 summarizes the results.

Table 2
 S parameters and corresponding bandwidth

S parameters boundary	Frequency [GHz]
BW ($S_{11} < -10$ dB)	12–18 (50%)
BW ($C = -20 \pm 2$ dB)	12–16.82 (40.16%)
$S_{14} < -25$ dB	12–17
$S_{12} > -0.5$	12–18

4. Multi-hole Coupler Design

The multi-hole technique is used to increase the directivity and bandwidth. To eliminate output power at port 4, the distances between the holes must be about $\lambda_g/4$, where λ_g is the guiding wavelength of the dominant TE_{10} mode at the center frequency [18]. In Fig. 6, the geometry of the designed four-hole Chebyshev coupler for Ku band is shown. Formulas from [18] were used to calculate the dimensions

of the multi-hole coupler for the coupling coefficient of -20 dB, with the thickness of coupling slots being considered to equal zero. 3D and 2D views of the coupler design are shown in Fig. 6a and Fig. 6b, respectively.

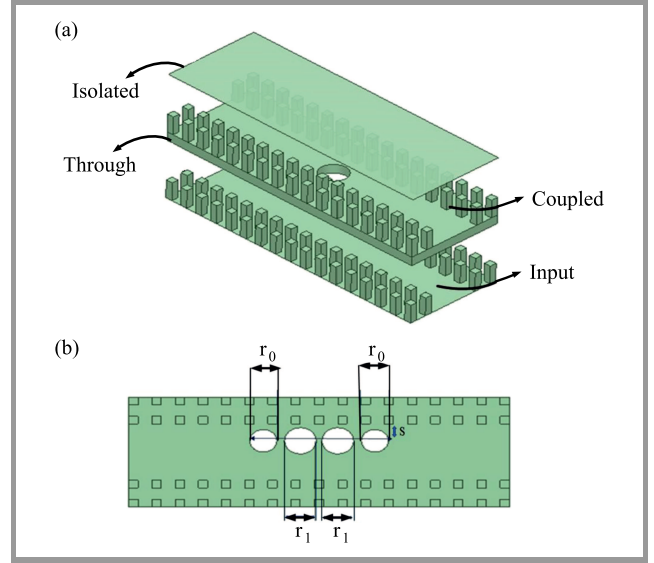


Fig. 6. Multi-hole GGW Bethe hole coupler: (a) 3D view, (b) top view.

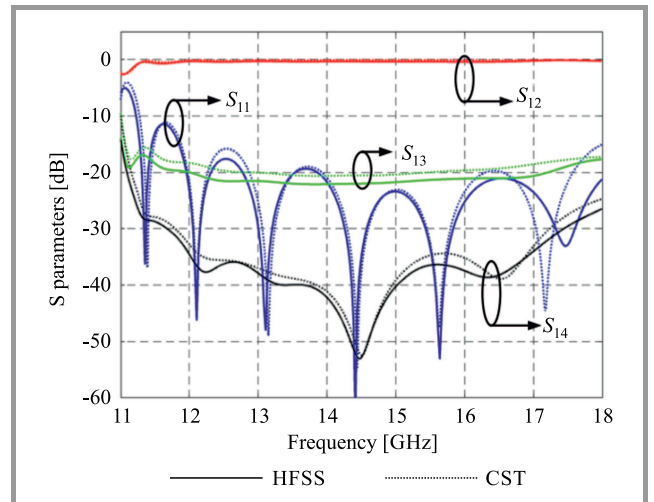


Fig. 7. S parameters of the designed 20 dB multi-hole coupler.

Similarly to the single-hole coupler, the effect that the thickness of slots has on the coupling coefficient are investigated at three frequencies within the operating bandwidth, and the same behavior as in the case of the single-hole coupler was observed. To achieve the desired coupling coefficient of -20 dB, the dimensions of the slots are fine-tuned: $r_0 = 2.8$ mm, $r_1 = 3.1$ mm, and $S = 3.75$ mm.

The S parameters for the coupling slot thickness of 1-mm are shown in Fig. 7 and in Table 3.

S_{11} is lower than -10 dB and S_{12} is more than -0.3 dB over the majority of the frequency band, evidencing proper impedance matching, as well as a low loss rate. The coupling coefficient S_{13} is almost -20 dB, which indicates

Table 3
Summary of simulation results

S parameters boundary	Frequency [GHz]
BW ($S_{11} < -10$ dB)	11–20 (81.81%)
BW ($C = -20 \pm 2$ dB)	11.22–17.86 (59.44%)
$S_{14} < -30$ dB	11–17.51
$S_{12} > -0.3$	11–20

a flat and proper coupling performance. S_{14} also remains below -30 dB, which confirms very good isolation between ports.

5. Comparison of Single-hole and Multi-hole Couplers

The results shown in Figs. 5 and 7 confirm the improvement in the structure's performance in terms of insertion loss (S_{12}), return loss (S_{11}), bandwidth and isolation, with the said improvement originating from the increasing number of holes. In both cases the coupling coefficient is very near to the desired value of -20 dB, and it remains flat over the entire bandwidth. Figure 8 compares the performance of the designed single- and multi-hole couplers in terms of directivity, showing that the multi-hole version provides higher directivity.

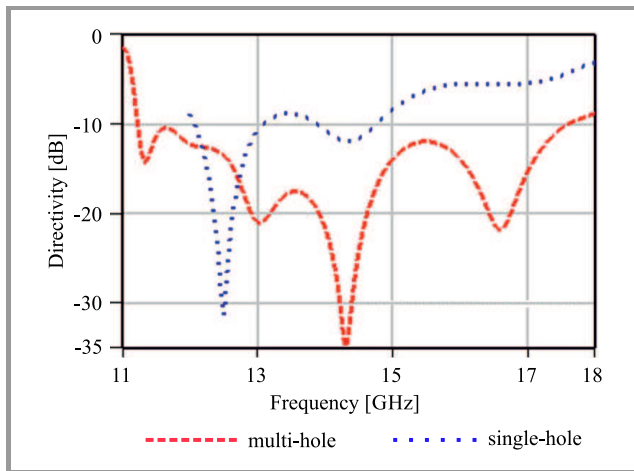


Fig. 8. Directivity of both structures.

6. Conclusions

In this paper, two types of 20 dB Bethe couplers are designed and simulated using groove gap waveguide technology. Simulation results show desirable performance of both designed structures. The couplers are suitable for Ku band operation. However, it is observed that the multi-hole version shows better performance in terms of insertion loss, return loss, bandwidth and directivity.

References

- [1] M. Bozzi, M. Pasian, L. Perregrini, and K. Wu, "On the losses in substrate integrated waveguides", in *Proc. Eur. Microwave Conf.*, Munich, Germany, 2007, pp. 384–387 (doi: 10.1109/EUMC.2007.4405207).
- [2] F. Xu and K. Wu, "Guided-wave and leakage characteristics of substrate integrated waveguide", *IEEE Trans. on Microw. Theory and Techniq.*, vol. 53, no. 1, pp. 66–72, 2005 (doi: 10.1109/TMTT.2004.839303).
- [3] M. Pasian, M. Bozzi, and L. Perregrini, "A formula for radiation loss in substrate integrated waveguide", *IEEE Trans. on Microwave Theory and Techniques*, vol. 62, no. 10, 2014 (doi: 10.1109/TMTT.2014.2341663).
- [4] J. Liu, A. U. Zaman, and J. Yang, "Design of wideband slot array antenna by groove gap waveguide in millimeter waves", in *Proc. IEEE-APS Topical Conf. on Antenn. and Propag. in Wirel. Commun. APWC 2018*, Cartagena des Indias, Colombia, 2018, pp. 725–728 (doi: 10.1109/APWC.2018.8503800).
- [5] A. Vosoogh *et al.*, "W-band low-profile monopulse slot array antenna based on gap waveguide corporate-feed network", *IEEE Trans. on Antenn. and Propag.*, vol. 66, no. 12, pp. 6997–7009, 2018 (doi: 10.1109/TAP.2018.2874427).
- [6] L. Wang *et al.*, "Low-dispersive leaky-wave antenna integrated in groove gap waveguide technology", *IEEE Trans. on Antenn. and Propag.*, vol. 66, no. 11, pp. 5727–5736, 2018 (doi: 10.1109/TAP.2018.2863115).
- [7] D. Zarifi, A. Farahbakhsh, A. U. Zaman, and P.-S. Kildal, "Design and fabrication of a high-gain 60-GHz corrugated slot antenna array with ridge gap waveguide distribution layer", *IEEE Trans. on Antenn. and Propag.*, vol. 64, no. 7, pp. 2905–2913, 2016 (doi: 10.1109/TAP.2016.2565682).
- [8] A. U. Zaman and P.-S. Kildal, "Different gap waveguide slot array configurations for mmwave fixed beam antenna application", in *Proc. 10th Eur. Conf. on Antenn. and Propag. EuCAP 2016*, Davos, Switzerland, 2016, pp. 1–4 (doi: 10.1109/EuCAP.2016.7481541).
- [9] A. Vosoogh *et al.*, "E-band 3-D metal printed wideband planar horn array antenna", in *Proc. In. Symp. on Antenn. and Propag. ISAP2016*, Okinawa, Japan, 2016, pp. 304–305.
- [10] R. Maaskant *et al.*, "Spatial power combining and splitting in gap waveguide technology", *IEEE Microw. and Wirel. Components Lett.*, vol. 26, no. 7, pp. 472–474, 2016 (doi: 10.1109/LMWC.2016.2574828).
- [11] M. S. Sorkherizi, A. Vosoogh, A. A. Kishk, and P.-S. Kildal, "Design of integrated diplexer-power divider", in *Proc. IEEE MTT-S Int. Microw. Symp. IMS 2016*, San Francisco, CA, USA, 2016 (doi: 10.1109/MWSYM.2016.7540124).
- [12] A. Vosoogh, A. A. Brazález, and P.-S. Kildal, "A V-band inverted microstrip gap waveguide end-coupled bandpass filter", *IEEE Microw. and Wirel. Compon. Lett.*, vol. 26, no. 4, pp. 261–263, 2016 (doi: 10.1109/LMWC.2016.2538598).
- [13] P.-S. Kildal *et al.*, "Design and experimental verification of ridge gap waveguide in bed of nails for parallel-plate mode suppression", *IET Microw., Antenn. and Propag.*, vol. 5, no. 3, pp. 262–270, 2011 (doi: 10.1049/iet-map.2010.0089).
- [14] E. Rajo-Iglesias and P.-S. Kildal, "Groove gap waveguide: A rectangular waveguide between contactless metal plates enabled by parallel-plate cut-off", in *Proc. of the 4th Eur. Conf. on Antenn. and Propag.*, Barcelona, Spain, 2010.
- [15] A. A. Brazalez, A. U. Zaman, and P.-S. Kildal, "Improved microstrip filters using PMC packaging by lid of nails", *IEEE Trans. on Compon., Packag. and Manufact. Technol.*, vol. 2, no. 7, pp. 1075–1084, 2012 (doi: 10.1109/TCPMT.2012.2190931).
- [16] A. Valero-Nogueira *et al.*, "Gap waveguides using a suspended strip on a bed of nails", *IEEE Antenn. and Wirel. Propag. Lett.*, vol. 10, pp. 1006–1009, 2011 (doi: 10.1109/LAWP.2011.2167591).
- [17] L. F. Carrera-Suárez *et al.*, "A novel twist between Gap Waveguides for compact slot-array antennas", in *Proc. IEEE Antenn. and Propag. Soc. Int. Symp. APSURSI 2014*, Memphis, TN, USA, 2014, pp. 456–457 (doi: 10.1109/APS.2014.6904559).

- [18] D. M. Pozar, *Microwave Engineering*, 4th ed. Wiley, 2011 (ISBN: 978-0-470-63155-3).
- [19] H. Bethe, "Theory of diffraction by small holes", *Phys. Rev.*, vol. 66, pp. 163–182, 1944 (doi: 10.1103/PhysRev.66.163).



Efat Nematpour received her B.Sc. degree in Electrical Engineering from the University of Lorestan, Iran. She works at the Department of Electrical and Computer Engineering, Graduate University of Advanced Technology, Kerman, Iran.

E-mail: Efatnematpour@yahoo.com
Graduate University of Advanced Technology
Kerman, Iran



Mohamad Hossein Ostovarzadeh received his B.Sc. degree in Electrical Engineering from Shahid Bahonar University, Kerman, Iran, in 2004, and M.Sc. degree in Telecommunications from the Sharif University of Technology, Teheran, Iran, in 2006. He received his

Ph.D. degree in Electrical Engineering from the Amirkabir University of Technology. His research interests are in the field of electromagnetic non-destructive testing techniques and microwave circuits.

E-mail: mh.ostovarzadeh@kgut.ac.ir
Graduate University of Advanced Technology
Kerman, Iran



Seyed Ali Razavi received his B.Sc. degree in Electrical Engineering from the University of Shahid Bahonar, Kerman, Iran, in 2006 and M.Sc. in Electrical Engineering from Ferdowsi University of Mashhad, Iran, in 2009. He received his Ph.D. degree in Electrical Engineering from Ferdowsi University of Mashhad, Iran, in 2013. Currently,

he works at the Graduate University of Advanced Technology, Kerman, Iran, as an Assistant Professor. He also collaborates with the group of antenna experts at the Chalmers University of Technology. His interests focus on microwave and millimeter-wave passive devices, antennas, electromagnetic wave scattering, SIW and gap-waveguide technologies.

E-mail: s.razaviparizi@kgut.ac.ir
Graduate University of Advanced Technology
Kerman, Iran

A Survey on Data Perception in Cognitive Internet of Things

Lokesh B. Bhajantri¹ and Prashant Baluragi²

¹ Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkot, India

² Department of Master of Computer Application, KLE Institute of Technology, Hubli, India

<https://doi.org/10.26636/jtit.2019.131419>

Abstract—A Cognitive Internet of Things (CIoT) is a brand of Internet of Things (IoT) with cognitive and agreeable mechanisms, which are incorporated to advance performance and accomplish insights into real world environments. CIoT can perceive present system's conditions, analyze the apparent information, make smart choices, and increase the network performance. In this survey paper, we present classifications of data perception techniques used in CIoT. This paper also compares the data perception works against energy consumption, network life-time, resource allocation, and throughput, as well as quality of data and delay. In addition, simulation tools for IoT and their performance are discussed. Finally, we provide the model of cognitive agent-based data perception in CIoT for future research and development, which ensures the network performance in terms of reliability, energy efficient, accuracy, scalable, fault tolerant, and quality of data.

Keywords—Internet of Things, Cognitive IT, data perception, cognitive agents.

1. Introduction

Minerva *et al.* in [1] described that Internet of Things (IoT) is a novel paradigm which enables many relative small devices like sensors/actuators, to acquire and transfer information to distributed frameworks without or with negligible human intervention. The advancement of IoT has brought an awesome impact to lot of applications in then allows to use technologies like social insurance, transportation, assembling, as discussed by Gubbi *et al.* and Luong *et al.* in [2], [3]. The IoT associates every one of the things in the physical world to the Internet, and cloud computing, deep learning, cognitive computing. The communication utilizing IoT for different applications is given in Fig. 1. This figure comprises with three layers for IoT communications is as follows:

- **Mobile-to-mobile (M2M) device.** A “thing” (in IoT) is an object equipped with sensors or/and actuators. Sensors gather data which next are transferred over a network. Actuators allow things to act (for example, to switch on or off the light. This concept includes the sensors in various domains like medical

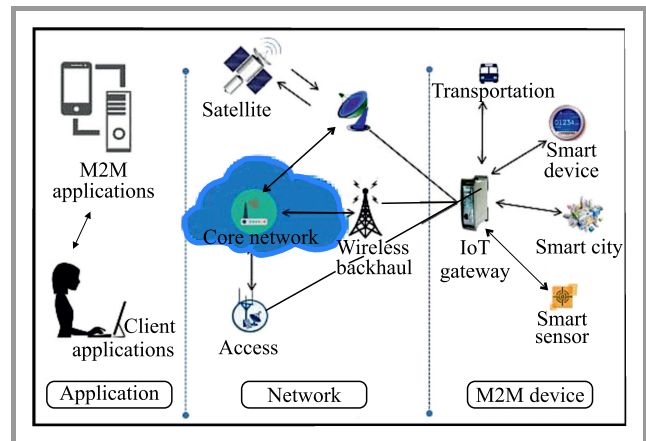


Fig. 1. Communications scheme in IoT.

and healthcare, public safety and environment monitoring, smart home/building, smart grid, smart mobility and smart tourism, industrial processing, agriculture, breeding, logistics, and product lifetime management.

- **Core network.** Specifies various access points (through wireless/wired) which transmits and receives data from gateways.
- **Applications.** Includes M2M and client applications. Applications are software components which acts as interface between an IoT system and users. These applications facilitate the users by giving options to monitor and control their devices (while they are connected to a network of similar devices, for example, homes or cars and controlled by a central system).

Therefore, the data sensed or gathered by various devices is accessed through wireless or wired network by gateways and connected to the end users of various applications where data is monitored and controlled.

Dynamic innovations in the fields from wireless sensors to nanotechnology are required for the development an IoT system. However, without comprehensive cognitive capability, IoT is just like an awkward stegosaurus, all brawn and no brains. To fulfill its potential and deal with growing

Table 1
Comparative analysis of the performance parameters employed in IoT applications

Author	Energy consumption	Life time	Resource allocation	Throughput	Response time	Quality of Service	Delay
Li <i>et al.</i> [4]	Low	–	Good	–	–	–	High
Hong <i>et al.</i> [5]	Low	–	–	–	Low	–	–
Baloch <i>et al.</i> [6]	Low	–	Poor	–	High	Better	–
Zhu <i>et al.</i> [7]	Low	–	–	High	–	Good	–
Yang <i>et al.</i> [8]	Low	–	Good	–	–	–	–
Bhardwaj <i>et al.</i> [9]	Low	–	Good	High	–	Good	Low
Chen <i>et al.</i> [10]	Low	Better	–	High	–	–	High
Naser <i>et al.</i> [11]	Low	–	–	–	Low	–	–
Feng <i>et al.</i> [12]	–	–	–	Low	–	Better	–

Table 2
Comparison of few platforms used to build various applications in IoT and CIoT

Platforms	Pros	Cons	Applications
NS2 Simulator	- relatively cheaper and can be used in testing complex scenarios - simulates across various protocols and platforms	- simulating real system cannot be relied - uncertainty in results statistics	can be used to demonstrate data traffic from sensors to host
Cooja Simulator	- can be used to visualize the status of each node - parameters like transmission range, radio interferences, propagation delays, receiver power can be obtained	comparatively slower emulation at the hardware level	- security key exchange between nodes can be demonstrated - helps evaluate interoperability among network stacks, energy consumption
Matlab tool	- remove noise and outliers - perform statistical analysis on channel data and explore trends	it is a paid software tool	works with ESP8266 chip and Raspberry Pie
Opportunistic Network Environment (ONE) simulator	- generates node motion using movement models - helps in visualizing both mobility and message passing in real time with a GUI	- no physical and MCA layer support - event can happen only before or after simulation	used in demonstrating DTN protocols to analyze delays in data transfer of no end-to-end paths communication
Java programming	- require less processing power - portable across various hardware platforms	- compiler is not well optimized - byte-code interpretation is slow	can be used to run embedded and small devices
C++	- capable of handling complex tasks with high speed of execution - hardware level programming is comparatively simpler	- no sophisticated garbage collection - strongly typed language aids in remembering all syntaxes	- used in developing resource constrained devices - useful for embedded system developers

challenges, we must take the cognitive capability into consideration and empower IoT with high-level intelligence, i.e., brain-empowered IoT or CIoT.

In this paper, we have analyzed some measures like energy consumption, network life time, resource allocation,

throughput, time duration, quality of service and delay researched by various authors (Table 1). Since their vital role in any network, we tried to consider and compared these parameters, which would give us the scope to consider their importance while designing and implementing/simulating

methodology for data perception in research work. Some of the parameters are considered for performance of IoT applications, these are shown in Table 1:

- **Energy consumption.** The energy is the major constraint resource of IoT/CIoT. The energy consumption is based on the amount of data perception or data processing or data gathering by the various nodes in the environment. Which will also depend on the applications used in the environment.
- **Network lifetime.** It is defined as the increase the probability of cognitive approaches in IoT as the increase the network lifetime of IoT. This is the major important metrics for the networks evaluation. The network lifetime parameter is also depending upon the efficient resources utilization in IoT/CIoT environment.
- **Resource allocation.** The mechanism for allocating resources (could be bandwidth, energy, nodes, etc.) allows to perform better network lifetime and increase the fault tolerance of IoT/CIoT.
- **Throughput.** It is the average rate of successful message delivery over a network (i.e. the number of data sent and received successfully over the network). It is an important representation of the performance and quality of a network connection.
- **Response time.** Response time is a measure of the total elapsed time between the start and completion of a task. All IoT/CIoT objects should be responding within the time constraints. There are two types of response time in IoT/CIoT like hard response time and soft response time. In hard response time, all IoT/CIoT objects should be responded to within the time constraints but should not violate the time constraints whereas soft response time, objects may violate the time constraints occasionally in the network environment.
- **Quality of Service.** QoS refers to any technology that manages data traffic to reduce packet loss, latency and jitter on the network. QoS controls and manages network resources by setting priorities for specific types of data.
- **Network delay.** The network delay specifies how long it takes for a bit of data to process across the network from one node to another.

In this paper, we have also identified few platforms used to build IoT applications, their pros and cons, and application (Table 2).

While the current development stage renders itself as a promising answer for enhancing availability, its reasonableness with regards to IoT is restricted for primary reasons are high cost, energy utilization, and lifetime. Currently, the most recognized attributes of IoT is that the

expanding between network among general things or items and various intriguing services or applications. In any case, so far huge numbers of the current IoT applications are as yet subordinate exceedingly on individuals for cognition processing. This perception fills in as one of the essential inspirations of this paper to present CIoT, where general objects behave as agents and associate with physical environment as well as social communities with least human intercession.

2. Cognitive IoT

IoT is a technological revolution that brings into a new ubiquitous connectivity, computing, and communication era as described by Qihui *et al.* [13]. The fundamental concept is to equip any physical objects with sensors and collaborate with Internet to synchronize the physical and digital world, which is together to achieve many ubiquitous and pervasive computing-based applications and further used for the automation. The second generation of IoT is known as Cognitive Internet of Things (CIoT). Currently, the data gathered by the sensors/objects are not utilized properly in terms of bandwidth, security, and privacy constraints as well as enormous unstructured data from logs and social frameworks are abundant in quantity of data. These data can be combined in a real time basis and valuable informations can be extracted by CIoT capabilities. The followings are the list of issues and challenges in CIoT mentioned by Dohr *et al.* [14]:

- standards for cross interoperability with heterogeneous networks,
- machine to machine standardization,
- self aware and self organizing networks,
- open framework for the CIoT,
- fault tolerance,
- adhoc deployable and configurable networks,
- dynamic and adaptable interoperability,
- open platform for CIoT,
- distributed energy efficient data processing,
- security and routing,
- access control and accounting schemes for IoT,
- low cost, secure, and high performance authentication,
- cognitive computing for IoT,
- context-aware computing,
- resource management for CIoT,
- connectivity,

- power,
- adaptability,
- data perception,
- scalability.

In addition, the knowledge of IoT has been accentuated to an ever-increasing extent. After an intensive examination on the models of IoT, Zhang *et al.* [15] additionally trusted that the knowledge still can't fulfill the need of use. Subsequently, another idea of CIoT is proposed through coordinating insight into IoT. A CIoT is an IoT with cognitive capacity which is incorporated to advance execution and accomplish insight is characterized by Kasnesis *et al.* [16]. CIoT can apperceive current business writes and organize conditions, break down the apparent data in light of the earlier learning, make intelligent choices, and perform versatile and control activities, which aims to expand network execution and meet the application necessity.

In our investigation, we have incorporated cognition into IoT and present another idea of CIoT. The idea of CIoT is to allocate the intelligence in IoT. Cognition can impenetrate any parts of IoT and goes about as an essential part for a given application with a specific end goal to accomplish above examined points. The network model of CIoT is presented and it consists of three imperative parts showed up in Fig. 2:

- **Access layer.** Specifies the data access through internet from various sensors and actuators equipped in many physical devices and data from different

domains like medical and healthcare, databases. Data analysts make use of such data to find trends and gain actionable insights.

- **Core network.** Consists of network instruction control channel where in gateway that involves in preprocessing and filtering before moving it to the access points (to reduce the volume of data for detailed processing and storing) and transmits control commands going from the Internet to access points and finally to IoT based applications.
- **Application layer.** Cognitive agents having high reasoning capabilities process the data from core network and resolves real time problems that have a high degree of dynamism which is very essential for data perception in IoT. The solution to real time problems is realized in the various user applications like health-care, traffic monitoring, and smart city/building.

The model incorporates core network and different access network spaces. The core network is principally comprised of access, remote, transmission routers, and so on. The access network spaces cognitive nodes, simple nodes, and different terminals. The implications of components are depicted in Fig. 2.

- **CIoT objects domains.** CIoT objects domains are high coupled and relatively independent. It is observed that, CIoT application consist at least one object domain but object domain can be divided into several sub-object domains. Based on the capabilities of data accessing nodes, these can be divided as cognitive nodes (CNs) and simple nodes (SNs). There are different numbers of CNs in different object domains and possible only one CN in an object domain under special circumstances. If object domain includes multiple CNs, two or more CNs can work together to achieve goals according to requirements. Furthermore, two or more object domains can also cooperate if necessary.
- **CIoT servers domain.** Object domains can apperceive the crude data, however without the ability to manage it. The data will be sent to server running unique schedules to be handled. For the most part, servers for an application are coupled and fault-tolerance, in this way, they are named as server domain. Now and then, server domain might be just a normal PC or a control terminal.
- **Cognitive nodes.** For multi domain cooperation (MDC), it alludes the particular CN to choose from every space to do agreeable assignments. There is diverse number of CNs in various spaces, possibly just a single. The demonstrating of CIoT design to locate new research thought for IoT and additionally an application which be acquainted with the proposed engineering and show that the work will have far expansive application prospect.

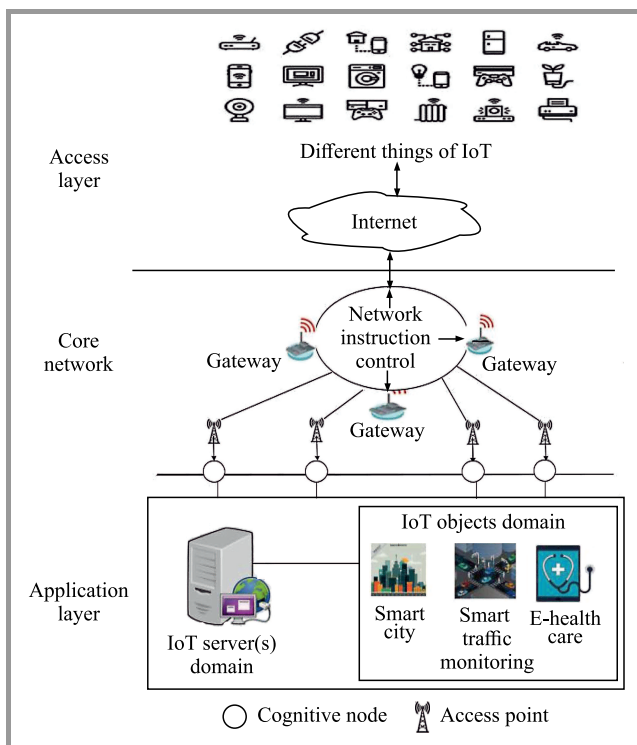


Fig. 2. CIoT network model.

Zhou *et al.* [17] has proposed a profound slant sentiment model inside the space of CIoT. Data recovery in social media is an overwhelming task which requires a great deal of specialized bits of knowledge. As a great instrument for large scale data recovery, hashing methods have likewise been widely utilized for content recovery. Deep sentiment hashing model is introduced, which is made out of following three stages are:

- a progressive consideration based long short-term memory network (LSTM) is prepared to acquire assessment particular report portrayals,
- given the record embeddings, k -nearest neighbor (kNN) calculation is utilized to develop a Laplacian matrix which is anticipated into hash marks by means of Laplacian eigenmaps (LapEig) later,
- model for hash functions learning, which is regulated by both the produced hash names and the first sentiment names.

Such joint supervision guarantees that a definitive hash codes delivered by the learned hash functions keep up sentiment level closeness.

Zhan *et al.* [18] have proposed the cloud assisted CIoT in view of the registering and information storage capacities of cloud virtual machines, which make ensuring virtual machine, file system critical for the entire framework security. Agent less intermittent file system screens are ideal solutions for ensure cloud virtual machines on account of the safe and low-overhead highlights. Notwithstanding, the greater part of the intermittent screens generally check the majority of the virtual machine file system or ensured documents in each scanning survey, so plenty of secure records are examined over and over despite the fact that they are not defiled. A novel agent less occasional file system screen system is introduced for virtual machines with various image formats to enhance the execution of agent less intermittent screens. Cognitive radio with context of IoT is concerning a joined plan of future Internet and cognitive radio with respect to IoT. It proposes an embedded and simplicity agreeable spectrum identifying arrangement, which has been probably completed. Additionally, Alberti *et al.* [19] exhibited a dispersed programming-controlled sector aware spectrum distinguishing configuration to collect and dismember the range utilizes information. Finally, the proposed technique is joined to a future Internet configuration called NovaGenesis.

A device level multimodal data correlation mining (DMDC) demonstrate is quickly made in light out of the CCA to change the data incorporate into a subspace and dismember the data relationship. The relationship of the device is obtained in perspective of the intensive of data association and the region information of the device. By then a heterogeneous clustering model (HDC) is proposed by utilizing the result of the association examination to portray the device.

Lin *et al.* [20] proposed a device clustering algorithm in perspective of multimodal data correlation for CIoT, which joins the components of multimodal data association separate with device clustering. Somov *et al.* [21] proposed to improve crude sensor information with context data for facilitate utilization in CIoT. To do as such, authors have created virtual portrayals of real sensors called virtual items. These CIoT objects create crude sensor estimations, as well as enhance those with setting data also. The approach included, real sensor network organization for traffic checking in the city. Specifically, the work done exhibited how to improve crude information and how to separate knowledge from the accessible information.

The remainder of this paper is organized as follows. Section 3 presents the data perception in CIoT, pros and cons of the methods used for data perception by various authors. Section 4 presents architecture of cognitive agent and few data management techniques used by cognitive agent in CIoT. Section 5 describes proposed methodology for data perception in CIoT and applications.

3. Data Perception in CIoT

The current investigation in CIoT is far from the acknowledgment of conceptualized applications. There are different open research issues and difficulties are into this space and one of the significant issues in CIoT is data perception. Data perception is a procedure to guarantee the information being detected by the sensors that have better visually as far as applications. It additionally guarantees that how well the information is recouped from the sensor network without having any redundancies. The accompanying is some one of a kind qualities of sensor information amid data perception like: heterogeneous, high dimension, context-aware information, and nonlinear reparability. Because of these blended qualities it is exceptionally testing to process enormous information in CIoT applications, which are coordinated various advances and protocols. The improvement of viable semantic advances and information revelation strategies is still in early stage. Thus, the information observation in CIoT is as yet an open issue.

The perception and access technique in perspective of IoT for hard assembling resources, computational assets, and insightful assets were analyzed by Tao *et al.* [22]. A five-layered structure resource skillful perception and access framework in context IoT is arranged and shown. The key progressions for sharp perception and access of different assets (i.e. hard gathering and computational resources, and adroit resources) are depicted. Authors proposed another thought of CIoT and its relating insightful plan. CIoT contains protocol plane, cognitive plane and control plane modules and incorporates some novel arranged models, for instance, layer tradition stack, 3-layer cognitive rings. In paper [23], a number of STS sub-regions are found that respond selectively to certain types of social input, organized along a posterior to anterior axis. Authors also identified the regions of overlapping response to multiple contrasts,

Table 3
The methods used for data perception by various authors

Author	Method	Pros	Cons
Benincas <i>et al.</i> [25]	Cognitive-based perception oriented situation awareness	A novel approach, in view of the idea of perception; situation awareness improves basic decision-making processes and gives capable results	Enhancement of the proposed approach for managing extremely versatile circumstances can be done
Cui <i>et al.</i> [26]	Sparse representation in cognitive based robotic perception	Intensively including learning algorithms to accomplish highly adaptive and intelligent behaviors, including actuation, sensing, perception and adaptive control	Better execution is required as far as perception
Pappas <i>et al.</i> [27]	Cognitive and affective perceptions with fuzzy-set qualitative comparative analysis	More interactive to target clients' cognitive and affective perceptions, and increment their expectation to purchase	Suggesting that distinctive sorts of perceptions may co-exist so as to enhance perception
Gray <i>et al.</i> [28]	A new approach in light of fuzzy logic with cognition; Fuzzy cognitive mapping (FCM) for stakeholder perceptions	Capable of evolving rapidly and improved decision making More accurate, less time consuming	Improvements can be made to refine and extend the utilizations of fuzzy logic method; Limited life time
Guirguis <i>et al.</i> [29]	A multi-hop routing for cognitive radio networks	Better shows in good put routing overhead, average size	Leads to a higher aggregate end-to-end delay
Araujo <i>et al.</i> [30]	Perception based assessment for multidimensional data visualizations	Provides beginnings for further research on the role of perception in multidimensional data projections	May expand this by investigating additional datasets, more projection techniques
Nitti <i>et al.</i> [31]	SIoT (Social Internet of Things) paradigm for the sensing of channel status to execute cognitive radio solutions	Enables faster and reliable information/service	Less iterations are expected

including regions responsive to both language and theory of mind.

A creative approach presented by Chai *et al.* [24] relies upon fuzzy logic to show the fundamental cognitions and individual conduct standards of people on foot as deduced from perception in demand to evaluate age and sexual introduction effect of walkers in convergence a signalized crosswalk. Fuzzy sets and guidelines are made to show the association in the midst of human bits of knowledge and decisions of a man by walking. Through changing the enlistment components of different age and sexual introduction social events, individual direct principles of individuals by walking are surveyed and considered. Remarkable in connection to most past examinations, both more prepared and more energetic walkers are seen to be less risk taking than grown-up individuals by walking. Additionally, tremendous sexual orientation differentiates is found only for insights of most dangerous conditions.

As a summary the pros and cons of the methods used for data perception other authors are shown in Table 3.

4. Cognitive Agents

Cognitive agent (CA) is software that exhibits the core skills of a human brain to perform tasks like learning, reasoning, understanding, grasping and acquire knowledge. CAs involves in resolving real time problems that have a high degree of dynamism, which is, essential for data perception in CIoT as shown in Fig. 3. The following functions of CAs agents are:

- **Beliefs.** The beliefs represent to the data about the condition of the environment that is refreshed suitably after each detecting activity. The beliefs can be seen as the informative part of the framework state. Beliefs generate choices/options for desires.
- **Desires.** The desires are the motivational condition of the framework. They have data about the destinations to be expert, i.e. what needs or adjustments are related with the different current destinations. The way that the agent has a desire does not imply that the agent will do it. The agent completes

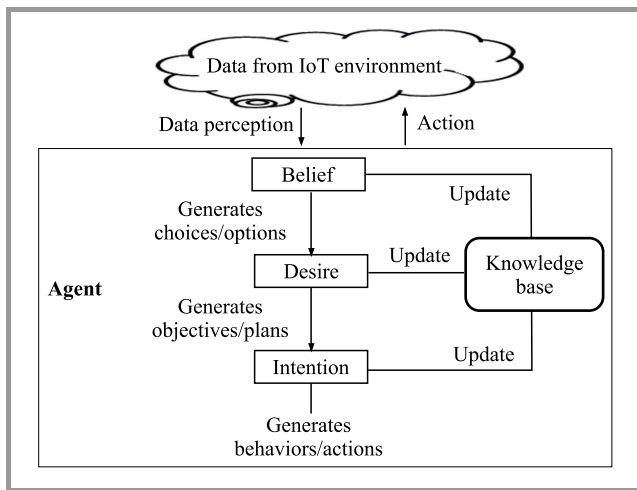


Fig. 3. Architecture of cognitive agents.

a deliberative procedure in which it stands up to its desires and beliefs and picks a set of desires that can be fulfilled. Desires generate objectives/plans to be implemented.

- **Intention.** The intention is a desire that is executed by a plan, since it can be done by the agents' beliefs (since it isn't rational that the agent completes something that it doesn't believe). Intentions generate actions as an output of this model architecture.
- **Knowledge base.** This is a repository that represents the agent's intentions on the basis of its current beliefs, desires, and intentions.

The desires can be conflicting to each other but not the intentions. The intentions represent the presently picked course of action. The intentions are insistent. An agent won't abandon its intentions – they will persist, until the point when the agent believes it has effectively accomplished them. In CIoT, the agent's techniques and behavior are depicted as the agent's beliefs. The choice of what to do and when to do it are the desires and intentions of the agent. Along these lines, a determined methodology (belief) of the agent is initiated if a desire of the agent turns into an intention. Followings are few data management techniques used by CAs in CIoT.

Bayesian networks. Perception is one of the required characteristics of cognitive process. The entire data that is perceived may not be important for cognitive process. Hence the classification of the perceived data is important and ideal, which is beneficial in saving working memory and ideal in long term memory. Thereby performance of the cognitive process can be enhanced as only relevant information is made available. The process of classification is carried out based on BDI model (belief, desire and intention) model using naive Bayes classifier is presented by Vishwanath *et al.* [32]. SWI-Prolog is used for the implementation purpose and results obtained, successfully clas-

sified perceived data as efficient and trivial perception based on BDI model using naive Bayes by first estimating the prior probability for effective and trivial perception and then likelihood probability. Anchisi *et al.* [33] have proposed that placebo analgesia as an aspect of pain perception, others being the modulating effects of emotions, cognition and past experience. Here authors suggest that computational considerations of pain may present a unifying justification of these phenomena. It is also shown how Bayesian decision theory can explain for such features which are described by a model of pain that is successfully tested against experimental data.

Fuzzy logic. Data perception using fuzzy cognitive mapping (FCM) has been used as a part of a different course of action of contexts ranging from fisheries organization to agricultural development, with a final objective to create clear graphical models of complex structures presented by Gray *et al.* [34]. This extension in reputation is an after effect of FCM's base up approach and its ability to merge an extent of individual, group level and master learning into an open and organized course of action. Regardless of the way that there has been an extension in the usage of FCM as ecological planning and learning device, limited progress has been made with regards to the methodology's relationship to existing adaptability structures and how the use of FCM contrasts and other participatory showing/approaches open. Li *et al.* [35] have composed a transmission scheduling technique named cognitive access for regular topology (CART) in cognitive radio sensor systems (CRSs) with insignificant control of channels and high transmission productivity. CART comprises of a timeslot and channel allotment method, an agreeable spectrum sensing plan, and a plan for detailing spectrum sensing results (SSRs). This paper likewise dissects CART's agreeable spectrum sensing for various transmission interference range and scope designs.

Neural networks. The multimedia neural cognitive computing (MNCC) model was designed by Liu *et al.* [36] based on the nervous mechanism and cognitive architecture. Furthermore, the semantic-oriented hierarchical cross-modal neural cognitive computing (CNCC) framework was proposed based on MNCC model and formal description and analysis for CNCC framework was given. It would effectively improve the performance of semantic processing for multimedia and cross-modal information. Zhang *et al.* [37] have proposed a unified deep learning based framework that enables effective human-thing cognitive interactivity in order to bridge individuals and IoT objects. They designed a reinforcement learning based selective attention mechanism (SAM) to discover the distinctive features from the input brain signals. In addition, authors proposed a modified long short-term memory (LSTM) to distinguish the inter-dimensional information forwarded from the SAM. Two practical real-time human-thing cognitive interaction applications were presented to validate the feasibility of their approach.

Context aware computing. Jamnal et al. [38] have proposed a new form of ambient intelligent architecture for smart homes by anticipating occupant's intentions for understanding purpose. This is achieved with a combined knowledge from traditional context aware pervasive systems and present epoch of CIoT technologies. Architecture of ambient cognition model (ACM) is prepared with artificial intelligence machine learning algorithms of subtractive clustering and hidden Markov model to identify the patterns in daily activity living. The main objective of ACM model is to provide comfortable living environment for occupant. This model can be successfully adopted in elderly health care systems, smart classrooms, and smart spaces for automated environment.

Knowledge structure tree. Cognitive model-based evolutionary mechanism was presented by Lu *et al.* [39] with the knowledge structure tree (KST). The information present in IoT was analyzed by mimicking the information processing system of human brain and the evolutionary mechanism of information status based on cognitive model was explained. The analyzed information can play a vital role for the implementation of intelligent information processing under the environment of IoT.

Clustering. Zhang et al. [40] have presented spectrum sensing method based on clustering algorithm and signal characteristic is proposed for cognitive wireless multimedia sensor networks. Initially the wireless communication signal characteristics were acquired according to the sampling signal covariance matrix. Then, the clustering algorithm was used to classify and test the signal characteristics. The results from the experiment conducted were sat-

isfactory and convinced about better sensing performance. Belk *et al.* [41] have focused on modeling users' cognitive approach on a set of web usage mining methods on navigation patterns and clickstream data. The purpose is to investigate whether k-means clustering can group users of particular cognitive approach by using measures acquired from a series of psychometric tests and content navigation activities. Three navigation metrics are proposed and used to find identifiable groups of users which have comparable navigation patterns in relation to their cognitive approach. The results of the experiment have been assessed with a user study which entailed a psychometric-based method for extracting the users' cognitive approach, combined with a real usage scenario of users navigating in a controlled web environment.

5. Proposed Methodology

The prime goal of the proposed study is to introduce a novel analytical modeling of data perception in CIoT for ensuring a robust, scalable, and cost-effective big data utilization. The paradigm of the proposed methodology is as follows (Fig. 4):

1. Deployment of wireless sensor network randomly.
2. Deployment of CAs in the WSNs.
3. Member nodes forward the context data to CH where CA is active.
4. CH after receiving data applies statistical measures to eliminate redundancies and forward the uniform data to gateway, which are further queued.

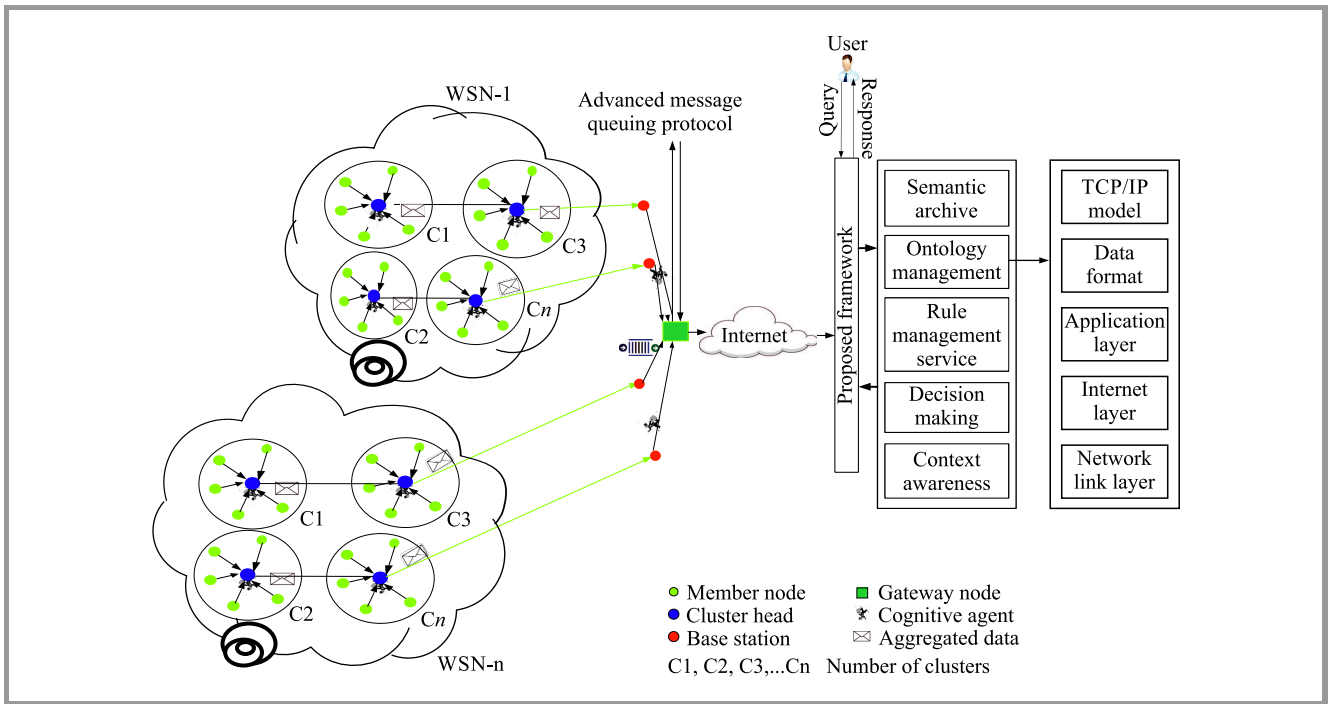


Fig. 4. Framework of data perception control using CA in CIoT. (For color pictures visit www.nit.eu/publications/journal-jtit)

5. Gateway implements advanced message queuing protocol to process all the incoming queues of data from CA.
6. The filtered aggregated data is then forwarded to the proposed framework via gateway where ontological operations are carried out followed by semantic operation.
7. The cognitive agent performs the rule management services to construct and impose rules (business-based rules) of data perception.
8. The processed data from rule management assists in decision making by using comparative analysis (incoming data and rule set) in CIoT.
9. Finally, the system model multiple contexts of data to further confirm the success factor of the rules (in order to avoid false positive).
10. Role of user is to query a particular request and receive specific and accurate response from the framework. The system is independent of any user-intervention.

Figure 5 illustrates the layout of proposed cognitive layered architectural model, the model consists of five layers: physical, data-perception, network, cognitive computing, and application layer.

ing, and application layer. The physical layer is associated with different heterogeneous (data sensed by various things/devices) WSNs. The access points are used as network administrators to access data sensed by physical layer. It's just not enough to collect the data from the access points but data should be addressed for various data characteristics like validity, completeness, accuracy, consistency, availability, and timeliness. Data perception layer consists of set of data functions like gathering, mapping, management, monitoring, merging, and semantic-ontology.

Cognitive nodes of network layer have cognitive computing features like:

- adaptive (ability to adjust to any surroundings),
- highly interactive (easy interaction with users),
- stateful (up-to-date),
- contextual (understand, identify the meaning).

Application layer includes different application domains like smart health care and monitoring, smart city/building, and so on. Finally, end user can access the CIoT applications. The objective of cognitive layer is to provide continuous services to users.

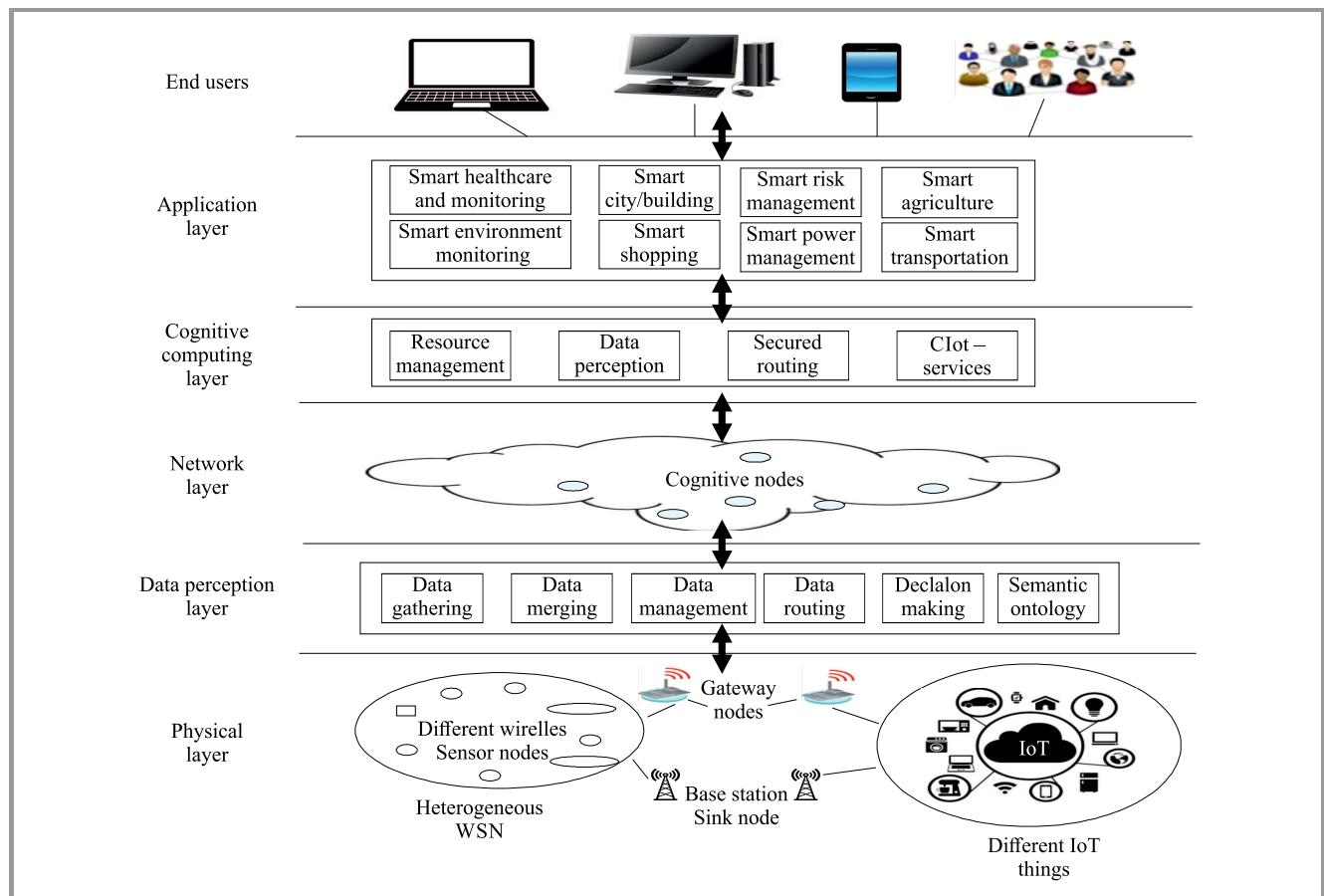


Fig. 5. Proposed layered architectural model in CIoT.

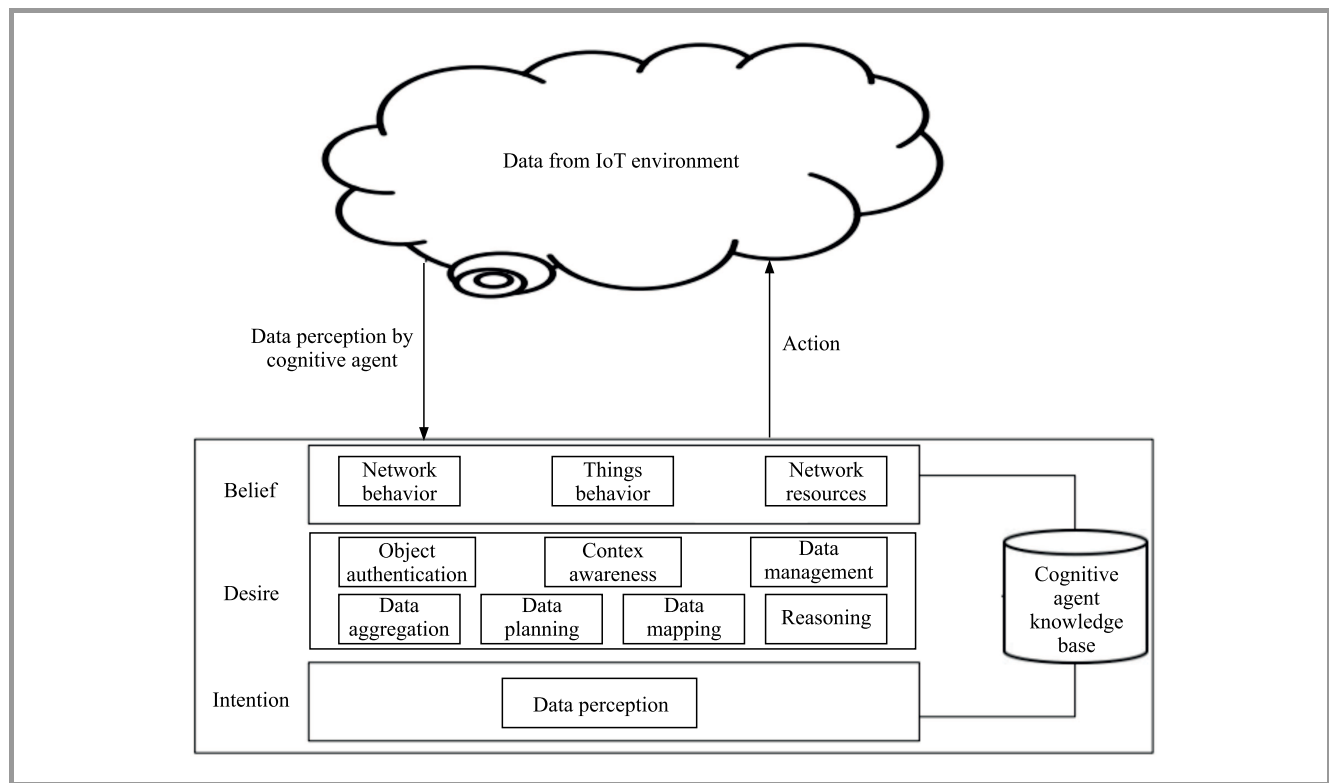


Fig. 6. Framework of cognitive agent BDI model for data perception.

5.1. Proposed Cognitive Agent BDI Model for Data Perception

The belief desire intention (BDI) model demonstration is given in Fig. 6. The cognitive agent constructs the belief model based on the data sensed or gathered by WSNs. CAs set the belief of the users based on the network behavior, things (from IoT) behavior and network resources. It is assumed that user is accessing the data from proper location and use the available network and devices. Based on the belief model, CA set the goals of the current environments, i.e. object authentication, context awareness, data management, data aggregation. Finally, agent defines the intentions of the environment which is to achieve data perception.

6. Conclusion

This paper describes the wide-ranging introduction of IoT and CIoT which covers issues and challenges and comparative analysis employed in IoT against the performance parameters like energy consumption, life time, resource allocation, throughput, time duration, quality of service, and delay. Also discussed CIoT environment reviewed on various pros and cons of different approaches for data perception in IoT environment. The review of cognitive agent based data perception is to present a new analytical demonstrating of data perception in CIoT for ensuring an effective use of big data. Presented the proposed methodology and model for data perception in CIoT using cognitive agents,

provide benefits like minimizing and simplifying the data process and minimizing handling of sensor data, ensure better quality of experience, better quality of information and more.

References

- [1] R. Minerva and A. B. Chebudie, "Towards a definition of the Internet of Things (IoT)", Technical Report, IEEE Internet Initiative, May 2015 [Online]. Available: https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf
- [2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions", *J. of Future Gener. Comp. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013 (doi: 10.1016/j.future.2013.01.010).
- [3] N. C. Luong *et al.*, "Data collection and wireless communication in Internet of Things (IoT) using economic analysis and pricing models: a survey", *IEEE Commun. Surv. & Tutor.*, vol. 18, no. 4, pp. 2546–2590, 2016 (doi: 10.1109/COMST.2016.2582841).
- [4] W. Li *et al.*, "Performance comparison of cognitive radio sensor networks for industrial IoT with different deployment patterns", *IEEE Systems J.*, vol. 11, no. 3, pp. 1456–1466, 2017 (doi: 10.1109/JSYST.2015.1500518).
- [5] G. J. Hong, S. Lee, J. Lim, W. Yoon, and S. Han, "RF spectrum sensing receiver system with improved frequency channel selectivity for Cognitive IoT Sensor Network applications", in *Proc. of IEEE MTT-S Int. Microw. Symp. IMS 2016*, San Francisco, CA, USA, 2016 (doi: 10.1109/MWSYM.2016.7540311).
- [6] Z. Baloch, F. K. Shaikh, and M. A. Unar, "A context aware data fusion approach for health-IoT", *Int. J. of Inform. Technol.*, vol. 10, no. 3, pp. 241–245, 2018 (doi:10.1007/s41870-018-0116-1).
- [7] J. Zhu, Y. Song, D. Jiang, and H. Song, "Multi-armed bandit channel access scheme with cognitive radio technology in wireless sensor networks for the Internet of Things", *J. on IEEE Access*, vol. 4, pp. 4609–4617, 2016 (doi: 10.1109/ACCESS.2016.2600633).

- [8] W. Yang and X. Zhao, "Robust resource allocation for orthogonal frequency division multiplexing based cooperative cognitive radio networks with imperfect channel state information", *J. of IET Commun.*, vol. 11, no. 2, pp. 273–281, 2016 (doi: 10.1049/iet-com.2016.0742).
- [9] P. Bhardwaj, A. Panwar, and O. Ozdemir, "Enhanced dynamic spectrum access in multiband cognitive radio networks via optimized resource allocation", *IEEE Trans. on Wirel. Commun.*, vol. 15, no. 12, pp. 8083–8106, 2016 (doi: 10.1109/TWC.2016.2612627).
- [10] Z. Chen, H. Wang, Y. Liu, F. Bu, and Z. Wei, "A context aware routing protocol on Internet of Things based on sea computing model", *J. of Computers*, vol. 7, no. 1, pp. 96–105, 2012 (doi: 10.4304/jcp.7.1.96-106).
- [11] M. Z. Naser and V. K. R. Kodur, "Cognitive Infrastructure – a modern concept for resilient performance under extreme event", *Autom. in Construction*, vol. 90, pp. 253–264, 2018 (doi: 10.1016/j.autcon.2018.03.004).
- [12] S. Feng, P. Setoodeh, and S. Haykin, "Smart home: cognitive interactive people-centric Internet of Things", *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 34–39, 2017 (doi: 10.1109/MCOM.2017.1600682CM).
- [13] W. Qihui *et al.*, "Cognitive Internet of Things: A new paradigm beyond connection", *IEEE J. on Internet of Things*, vol. 1, no. 2, pp. 129–143, 2014 (doi: 10.1109/JIOT.2014.2311513).
- [14] A. Dohr, R. Modre-Oprian, M. Drobics, D. Hayn, and G. Schreier, "The Internet of Things for ambient assisted living", in *Proc. of 7th Int. Conf. on Inform. Technol.: New Generation*, Las Vegas, NV, USA, 2010, pp. 804–809 (doi: 10.1109/ITNG.2010.104).
- [15] M. Zhang *et al.*, "A novel architecture for cognitive Internet of Things", *Int. J. of Secur. and its Appl.*, vol. 9, no. 9, pp. 235–252, 2015 (doi: 10.14259/ijisa.2015.9.9.21).
- [16] P. Kasnesis, C. Z. Patrikakis, D. Kogias, L. Toumanidis, and I. S. Venieris, "Cognitive Friendship and Goal Management for the Social IoT", *Computers and Elec. Engin.*, vol. 58, pp. 412–428, 2017 (doi: 10.1016/j.compeleceng.2016.09.024).
- [17] K. Zhou, J. Zeng, Y. Liu, and F. Zou, "Deep sentiment hashing for text retrieval in social CIoT", *J. of Future Gener. Comp. Syst.*, vol. 28, pp. 362–371, 2018 (doi: 10.1016/j.future.2018.03.047).
- [18] D. Zhan *et al.*, "A high-performance virtual machine file system monitor in cloud-assisted cognitive IoT", *J. of Future Gener. Comp. Syst.*, vol. 88, pp. 209–219, 2018 (doi: 10.1016/j.future.2018.05.055).
- [19] A. M. Alberti *et al.*, "Cognitive radio in the context of Internet of Things using a novel future internet architecture called NovaGenesis", *J. of Computers and Elec. Engin.*, vol. 57, pp. 147–161, 2017 (doi: 10.1016/j.compeleceng.2016.07.008).
- [20] K. Lin, D. Wang, F. Xia, and H. Ge, "Device clustering algorithm based on multimodal data correlation in Cognitive Internet of Things", *IEEE Internet of Things J.*, vol. 5, no. 4, pp. 2263–2271, 2017 (doi: 10.1109/JIOT.2017.2728705).
- [21] A. Somov, C. Dupont, and R. Giaffreda, "Supporting smart city mobility with Cognitive Internet of Things", in *Proc. Future Network & Mobile Summit*, Lisboa, Portugal, 2013 (ISBN: 978-1-905824-37-3).
- [22] F. Tao, Y. Zuo, L. D. Xu, and L. Zhang, "IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing", *IEEE Trans. on Indust. Informat.*, vol. 10, no. 2, pp. 1547–1557, 2014 (doi: 10.1109/TII.2014.2306397).
- [23] B. Deen, K. Koldewyn, N. Kanwisher, and R. Saxe, "Functional organization of social perception and cognition in the superior temporal sulcus", *Cerebral Cortex*, vol. 25, no. 11, pp. 4596–4609, 2015 (doi: 10.1093/cercor/bhv111).
- [24] C. Chai, X. Shi, Y. D. Wong, M. J. Er, and E. T. Meng, "Fuzzy logic-based observation and evaluation of pedestrians behavioral patterns by age and gender", *Transportat. Res. Part F: Traffic Psychol. and Behav.*, vol. 40, pp. 104–118, 2016 (doi: 10.1016/j.trf.2016.04.004).
- [25] G. Benincasa, G. D'Aniello, C. De Maio, V. Loia, and F. Orciuoli, "Towards perception-oriented situation awareness systems", in *Proc. of the 7th IEEE Int. Conf. Intelligent Syst. IS'2014*, Warsaw, Poland, 2014, pp. 813–824 (doi: 10.007/978-3-319-11313-5.71).
- [26] Z. Cui, S. S. Ge, Z. Cao, J. Yang, and H. Ren, "Analysis of different sparsity methods in constrained RBM for sparse representation in cognitive robotic perception", *J. of Intell. and Robotic Syst.*, vol. 80, no. 1, pp. 121–132, 2015 (doi: 10.1007/s10846-015-0213-3).
- [27] I. O. Pappas, P. E. Kourouthanassis, M. N. Giannakos, and V. Christikopoulos, "Explaining online shopping behavior with fsQCA: the role of cognitive and affective perceptions", *J. of Business Res.*, vol. 69, no. 2, pp. 794–803, 2016 (doi: 10.1016/j.jbusres.2015.07.010).
- [28] S. R. Gray *et al.*, "Are coastal managers detecting the problem? Assessing stakeholder perception of climate vulnerability using fuzzy cognitive mapping", *Ocean & Coastal Manag.*, vol. 94, pp. 74–89, 2014 (doi: 10.1016/j.ocecoaman.2013.11.008).
- [29] A. Guirguis, M. Karmoose, K. Habak, M. El-Nainay, and M. Youssef, "Cooperation-based multi-hop routing protocol for cognitive radio networks", *J. of Netw. and Comp. Appl.*, vol. 110, pp. 27–42, 2018 (doi: 10.1016/j.jnca.2018.03.005).
- [30] H. D. S. Araújo *et al.*, "A proposal for iot dynamic routes selection based on contextual information", *Sensors*, vol. 18, no. 2, pp. 353–344, 2018 (doi: 10.3390/s18020353).
- [31] M. Nitti, M. Murrioni, M. Fadda, and L. Atzori, "Exploiting social Internet of Things features in cognitive radio", *IEEE Access*, vol. 4, pp. 9204–9212, 2016 (doi: 10.1109/ACCESS.2016.2645979).
- [32] Y. Vishwanath, T. S. Murali, and M. V. Vijayakumar, "Implementation of perception classification based on BDI model using Bayesian classifier", in *Proc. of Int. J. of Comp. Sci. and Inform. Technol. (IJCSIT)*, vol. 5, no. 6, pp. 8161–8165, 2014 [Online]. Available: <http://ijcsit.com/docs/Volume%205/vol5issue06/ijcsit20140506282.pdf>
- [33] D. Anchisi and M. Zanon, "A Bayesian perspective on sensory and cognitive integration in pain perception and placebo analgesia", *PLoS One*, vol. 10, no. 2, pp. 1–12, 2015 (doi: 10.1371/journal.pone.0117270).
- [34] S. A. Gray *et al.*, "Using fuzzy cognitive mapping as a participatory approach to analyze change, preferred states, and perceived resilience of social-ecological systems", *Ecology and Society*, vol. 20, no. 2, pp. 1–10, 2015 (doi: 10.5751/ES-07396-200211).
- [35] L. Wenxiang *et al.*, "Performance comparison of cognitive radio sensor networks for industrial IoT with different deployment patterns", *IEEE Systems J.*, vol. 11, no. 3, pp. 1456–1466, 2017 (doi: 10.1109/JSYST.2015.2500518).
- [36] Y. Liu, "Research on the Brain-inspired Cross-media Neural Cognitive Computing Framework", 2018, arXiv:1805.01385 [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1805/1805.01385.pdf>
- [37] X. Zhang, L. Yao, S. Zhang, S. S., Kanhere, Q. Z., Sheng, Y. Liu, "Internet of Things meets brain-computer interface: a unified deep learning framework for enabling human-thing cognitive interactivity", *J. of IEEE Internet of Things*, vol. 6, no. 2, pp. 2084–2092, 2018 (doi: 10.1109/JIOT.2018.2877786).
- [38] G. Jamnal and X. D. Liu, "A cognitive-IoE approach to ambient intelligent smart home", in *Proc. of the 2nd Int. Conf. on Internet of Things, Big Data and Secur. IoTBDs 2017*, Porto, Portugal, 2017, vol. 2, pp. 302–308 (doi: 10.5220/0006304103020308).
- [39] L. Huimin, L. Hu, G. Liu, and J. Zhou, "Cognitive Model-based evolution mechanism of the information status in Internet of Things", *Inform. Technol. J.*, vol. 13, no. 15, pp. 2431–2436, 2014 (doi: 10.3923/itj.2014.2431.2436).
- [40] Y. Zhang, P. W. S. Zhang, Y. Wang, and N. Li, "A spectrum sensing method based on signal feature and clustering algorithm in cognitive wireless multimedia sensor networks", *Adv. in Multimed.*, vol. 2017, pp. 1–10, 2017 (doi: 10.1155/2017/2895680).
- [41] M. Belk, E. Papatheocharous, P. Germanakos, and G. Samaras, "Investigating the relation between users cognitive style and web navigation behavior with K-means clustering", in *Advances in Conceptual Modeling ER 2012 Workshops CMS, ECDM-NoCoDA, MoDIC, MORE-BI, RIGiM, SeCoGIS, WISM, Florence, Italy, October 15-18, 2012. Proceedings*, S. Castano, P. Vassiliadis, L. V. Lakshmanan, and M. L. Lee, Eds. LNCS, vol. 7518. Springer, 2012, pp. 337–346 (doi: 10.1007/978-3-642-33999-8.40).



Lokesh B. Bhajantri received Ph.D. degree in Computer Science and Engineering from the Visveswaraiyah Technological University (VTU), Belgaum, Karnataka, 2015. He is presently working as an Associate Professor in the Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkot,

India. His areas of interests include distributed sensor networks, cognitive Internet of Things, mobile computing and communications, networking protocols, genetic algorithms, applications of agents and real time systems. He is a member of Board of Studies in the Department of Information Science and Engineering, Basaveshwar Engineering College, Bagalkot, Karnataka, India. He is a member of International Association of Computer Science and Information Technology (IACSIT).

E-mail: lokeshcse@yahoo.co.in

Department of Information Science and Engineering
Basaveshwar Engineering College
Bagalkot, Karnataka, India



Prashant M. Baluragi received his M.Tech. degree in Computer Science and Engineering from the Basaveshwar Engineering College, Bagalkot under Visveswaraya Technological University Belagavi, India. He is presently, working as an Assistant Professor in Department of Master of Computer Applications, KLE Institute of

Technology, Hubballi, Karnataka, India. He has an experience of 5 years of teaching and an experience of 8 years in software industries (IBM India and TCS India). He is pursuing Ph.D. in the area of Cognitive Internet of Things (CIoT).

E-mail: prashant.baluragi@gmail.com

Department of Master of Computer Application
KLE Institute of Technology
Hubli, Karnataka, India

Information for Authors

Journal of Telecommunications and Information Technology (JTIT) is published quarterly. It comprises original contributions, dealing with a wide range of topics related to telecommunications and information technology. **All papers are subject to peer review.** Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

Suitable research-related papers should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published only by invitation.

Manuscript. TEX and LATEX are preferable, standard Microsoft Word format (.doc) is acceptable. The authors JTIT LATEX style file is available:

<http://www.nit.eu/for-authors>

Papers published should contain up to 10 printed pages in LATEX authors style (Word processor one printed page corresponds approximately to 6000 characters).

The manuscript should include an abstract about 150200 words long and the relevant keywords. The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results. Abstracts must not include mathematical expressions or bibliographic references.

Keywords should not repeat the title of the manuscript. About four keywords or phrases in alphabetical order should be used, separated by commas.

The original files accompanied with pdf file should be submitted by e-mail: redakcja@itl.waw.pl

Figures, tables and photographs. Original figures should be submitted. Drawings in Corel Draw and PostScript formats are preferred. Figure captions should be placed below the figures and can not be included as a part of the figure. Each figure should be submitted as a separated graphic file, in .cdr, .eps, .ps, .png or .tif format. Tables and figures should be numbered consecutively with Arabic numerals.

Each photograph with minimum 300 dpi resolution should be delivered in electronic formats (TIFF, JPG or PNG) as a separated file.

References. All references should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. Samples of correct formats for various types of references are presented below:

- [1] Y. Namihiro, Relationship between nonlinear effective area and mode field diameter for dispersion shifted fibres, *Electron. Lett.*, vol. 30, no. 3, pp. 262264, 1994.
- [2] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.
- [3] S. Demri and E. Orłowska, Informational representability: Abstract models versus concrete models, in *Fuzzy Sets, Logics and Knowledge-Based Reasoning*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301314.

Biographies and photographs of authors. A brief professional authors biography of up to 200 words and a photo of each author should be included with the manuscript.

Galley proofs. Authors should return proofs as a list of corrections as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within one week after receiving the offprint.

Copyright. Manuscript submitted to JTIT should not be published or simultaneously submitted for publication elsewhere. By submitting a manuscript, the author(s) agree to automatically transfer the copyright for their article to the publisher, if and when the article is accepted for publication. The copyright comprises the exclusive rights to reproduce and distribute the article, including reprints and all translation rights. No part of the present JTIT should not be reproduced in any form nor transmitted or translated into a machine language without prior written consent of the publisher.

For copyright form see: <http://www.nit.eu/for-authors>

A copy of the JTIT is provided to each author of paper published.

(Contents Continued from Front Cover)

A Modified Approach to Punctured Product Polar Codes

M. A. Hanif and S. Vafi

Paper

63

Ku Band Bethe Hole Coupler Using Gap Waveguide Technology

E. Nematpour, M. H. Ostovarzadeh, and S. A. Razavi

Paper

70

A Survey on Data Perception in Cognitive Internet of Things

L. B. Bhajantri and P. Baluragi

Paper

75

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

tel. +48 22 512 81 83
fax: +48 22 512 84 00
e-mail: redakcja@itl.waw.pl
<http://www.nit.eu>