Deep Reinforcement Learning with Factor Graphs for Optimized Path Selection in Full Duplex Wireless Multihop Networks

V Venkataiah, Kotikala Narayana Rao, Gude Ramarao, K Manoz kumar Reddy, Smitha Chowdary Chaparala, Naresh kumar Bhagavatham

***Abstract*—A wireless multihop network (WMN) uses wireless nodes to relay data packets without a controller. In such networks, source nodes' unpredictable routing decisions during multihop transmission might decrease network performance and throughput. To solve this, we present two new path selection algorithms: Noise-Level Path Selection (NLPS) and Interference-Level Noise-Learning Path Selection. Both approaches dynamically find the most effective multihop routes using deep reinforcement learning (DRL) to improve network end-to-end (E2E) throughput. To simplify processing, we use a nested lattice code (NLC) structure and a factor graph (FG) framework. Simulations show that NLPS and INLPS enhance network capacity by 3.1× and 10.5×, respectively, compared to traditional FG-based methods. NLPS takes 0.627 seconds, INLPS 1.221 seconds, and FG 0.006 seconds. Both strategies improve capacity and are feasible for real-time applications despite the longer processing time.**

***Index Terms*—Deep reinforcement learning, Factor graph, Nested lattice code, Learning path algorithm, Wireless multihop networks.**

I. Introduction

RECENT projections estimate that nearly 65% of the global population will be utilizing 5G networks in the near future [1]. This growing adoption signals a major increase in the use of connected devices, such as smartphones, smartwatches, and other mobile technologies. With the arrival of 6G, users can expect substantial enhancements over 5G, including extended coverage, higher data rates, support for massive connectivity, and ultra-low latency. These improvements will be driven by emerging technologies such as artificial intelligence (AI)-powered tools for real-time network management, planning, and optimization, along with extremely high-frequency bands (beyond 1 THz). Furthermore, 6G is anticipated to usher in a transformative digital revolution, transitioning from the traditional Internet of Things (IoT)" to the advanced concept of an Internet of Intelligence [1]. To meet this vision, 6G systems must deliver intelligent and ubiquitous AI-driven services across a wide spectrum of devices, from cloud infrastructures to edge terminals. AI will play a key role in designing and optimizing new architectures, communication protocols, and decision-making processes. Advanced technologies such as full-duplex communication, reconfigurable intelligent surfaces (RIS), and large-scale distributed MIMO systems are expected to be integral to future 6G systems. According to Trivedi et al. [2], wireless multihop networks (WMNs) will be a fundamental component of 6G, owing to their self-organizing capabilities, high scalability, decentralized architecture, and adaptability to dynamic environments.

In WMNs, nodes collaborate to autonomously establish a dynamic network structure, enabling data to travel over extended distances through multiple intermediary nodes. While these networks significantly extend coverage, they also introduce challenges such as reduced throughput and increased latency, particularly when inefficient routing paths are selected. As network size increases, so does the complexity of identifying the optimal transmission path among the many possible routes. This decision-making process directly influences network performance, as each route has unique characteristics and constraints. Additionally, the concurrent forwarding of packets—both local and relayed—by each node leads to delays due to processing and queueing, which become more severe with larger node counts. To address these challenges, our study proposes a solution that integrates deep reinforcement learning (DRL) with factor graph (FG) modeling and nested lattice code (NLC) strategies. The FG-based DRL (fDRL) framework is designed to enable each source node to select the most efficient relay node, forming a path that optimizes network throughput. FG assists in identifying the optimal root node for the network, effectively reducing computational overhead in DRL's iterative processes. Meanwhile, NLC, integrated with the compute-and-forward

Manuscript received April 24, 2025; revised July 5, 2025.

V Venkataiah is an Associate Professor of Computer Science and Engineering Department, CMR College of Engineering & Technology, Hyderabad, Telangana, India (E-mail: venkat.vaadaala@gmail.com).

K Narayana Rao is an Associate Professor of Computer Science and Engineering Department, RISE Krishna Sai Prakasam Group of Institutions, Ongole, Andhra Pradesh, India (e-mail: narayanarao2@gmail.com).

Gude Ramarao is an Associate Professor of Electronics and Communication Engineering Department, G. Pullaiah College of Engineering and Technology, Kurnool, Andhra Pradesh, India (e-mail: ramaraog19@gmail.com).

K Manoz Kumar Reddy is an Assistant Professor and Dean of Electrical and Electronics Engineering Department, Aditya University, Surampalem, Andhra Pradesh, India (e-mail: kmkreddyy@gmail.com).

Ch Smitha Chowdary is a Professor of Computer Science and Engineering Department, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India (e-mail: smithacsc@gmail.com).

Naresh kumar Bhagavatham is an Associate Professor of Computer Science and Engineering Department, Vignana Bharathi Institute of Technology, Ghatkesar, Hyderabad, Telangana, India (e-mail: bhagavatham.nareshkumar@vbithyd.ac.in).

(CoF) strategy, mitigates link errors and improves reliability, thereby enhancing network capacity by reducing the required transmission time slots.

Within this fDRL framework, we introduce two novel learning path selection (LPS) algorithms. The first leverages the signal-to-noise ratio (SNR) as a reward to boost end-to-end (E2E) throughput. The second extends this approach by employing the signal-to-interference-and-noise ratio (SINR) to further refine the selection of optimal paths for each source-destination pair. Numerous studies have explored techniques to improve WMN capacity, often focusing on analytical modeling, routing strategies, and power control mechanisms. For instance, Fujimura et al. [3] provided analytical insights into E2E throughput concerning hop count and payload size in linear topologies. S. Rezaei et al. [4] expanded this by integrating node distribution, routing strategies, and MAC layer considerations. Other researchers, such as Lee et al. [5] and Gui et al. [6], have proposed routing strategies that optimize for fairness and energy efficiency, respectively. Tree-based network structures have also gained attention, with Eliiyi et al. [7] proposing a parallel method to identify root nodes, thereby minimizing energy consumption and processing time. Yu et al. [8] introduced the CTPC technique to enhance E2E performance in densely populated WMNs, incorporating full-duplex communication for better results. Similarly, Khun et al. [9], [10] developed the OATC scheme to support concurrent transmission (CT) modes while maintaining low interference levels and high transmission rates regardless of network density.

Despite the promising applications of FG in wireless communications, its use in capacity modeling remains limited. For instance, Mao et al. [11] utilized FG in a centralized sensor network to iteratively estimate link loss, while Li et al. [12] presented a sequential particle-based SPA framework for distributed target state estimation in WSNs. Jiang et al. [13] applied FG to real-world scenarios like pedestrian tracking using smartphone-based dead reckoning. In the domain of coding theory, Bu et al. [14] incorporated network coding into WMNs to solve cooperative communication problems, optimizing scheduling, routing, and node selection simultaneously. However, lattice coding theory remains underexplored in WMNs. Xue et al. [15] demonstrated that lattice decoders can maintain a low word error rate even under stringent power constraints in wireless systems.

The growing role of AI in wireless communication has spurred a wave of research into AI-augmented network solutions. Rosenberger et al. [16] proposed a multi-agent DRL system for the Industrial IoT, allowing decentralized resource allocation with minimal computation delay. Inspired by this, our work adopts DRL to reduce path computation time in dynamic network environments. Wang et al. [17] presented a multigranular DRL strategy for channel allocation in WMNs, integrating mobile-edge computing to ensure reliable data transmission. Ho et al. [18] applied DRL in multi-access edge computing (MEC) to optimize offloading, server selection, and handovers in 5G networks, significantly enhancing performance and reducing overhead. Deep learning (DL), a prominent branch of AI, has also contributed to wireless network advancements [19], [20]. Reinforcement learning (RL), in particular, has shown promise for path selection tasks in WMNs. Dugaev et al. [21] introduced an adaptive packet-forwarding protocol based on RL, outperforming conventional routing in terms of reliability and recovery speed. Among RL techniques, Q-learning is widely applied due to its model-free nature and ability to function without predefined environment models. Wongphatcharatham et al. [22] demonstrated how multi-agent Q-learning can improve SINR by optimizing interference channels more effectively than traditional approaches. Similarly, Wang et al. [23] designed a Q-learning-based path selection method in clustered multihop networks, achieving near-optimal E2E throughput by decentralizing path computations. Su et al. \[24] proposed a Deep-Q-network solution to improve node selection in cooperative wireless sensor networks, significantly boosting both capacity and efficiency. Despite these advancements, no prior studies have investigated the integration of FG and NLC within a Q-learning-based routing framework. Therefore, our research uniquely contributes to this domain by merging FG modeling, Q-learning algorithms, and NLC strategies to significantly enhance network capacity and efficiency in wireless multihop network environments.

II. System model

This section introduces and elaborates on the proposed factor graph-based Deep Reinforcement Learning (fDRL) framework. The overall architecture of the fDRL system is illustrated in Figure 1. Within the context of a Wireless Mesh Network (WMN), numerous wirelessly connected devices—such as smartphones, robots, vehicles, and computers—are interconnected through multihop communication. The fDRL framework is composed of three primary components that collaboratively enable efficient information routing among the devices. First, in the root selection phase, a Factor Graph (FG) model is utilized alongside the Average Link Metric (ALM) and a Shortest Path Spanning Tree (SPST) algorithm to identify the most suitable root node within the tree-based network structure. Next, during the learning phase, two Q-learning-based algorithms—Naïve Learning Path Selection (NLPS) and Improved NLPS (INLPS)—are deployed to determine optimal forwarding paths for data packets from each device. In the final phase, known as the enhancement stage, the Node-Link Clustering (NLC) and Coordination Function (CoF) are employed to manage and optimize data transmissions between paired nodes.

This strategy enhances communication reliability, especially in environments with high noise or interference, and reduces the required time slots for data delivery. This approach is particularly well-suited for large-area network deployments, such as those found in stadiums or exhibition centers, where communication devices are largely stationary and typically rely on a single-root tree topology for data exchange. The Factor Graph (FG) model, in this context, is a bipartite graphical representation that decomposes a global function into a product of smaller, local functions. FG comprises two main types of nodes: variable nodes and factor nodes.

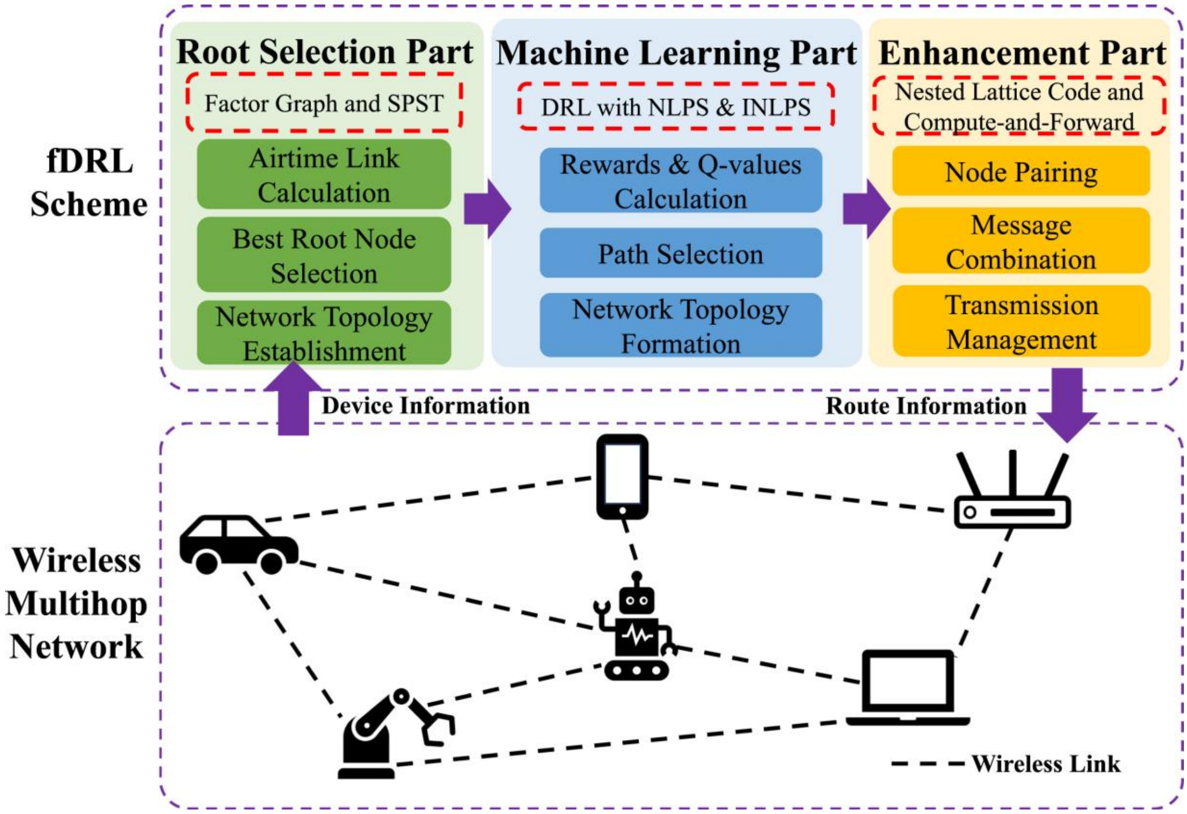


Fig. 1. The fDRL scheme's structural diagram

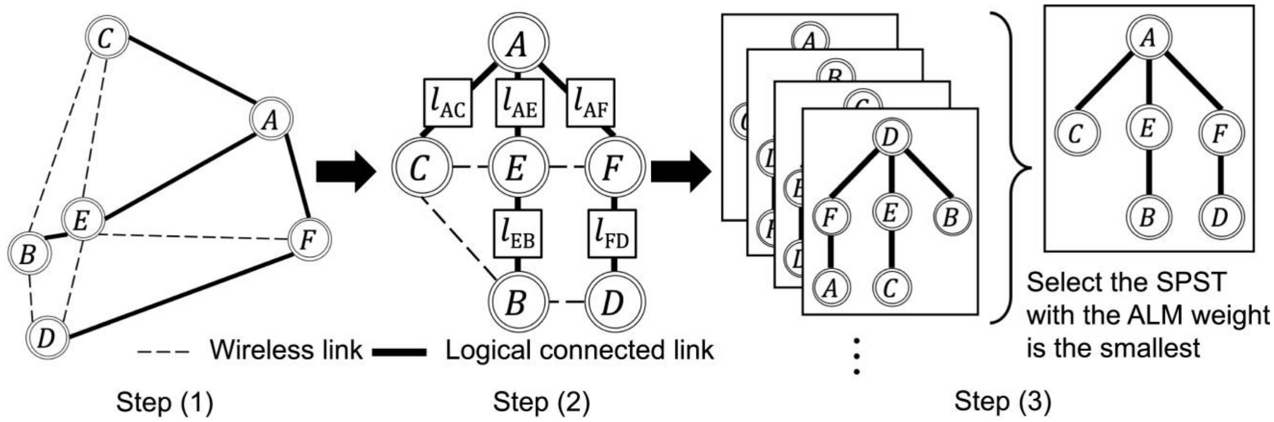


Fig. 2. SPST calculation for node A utilizing Dijkstra's algorithm

Variable nodes represent known values (evidence variables) or values to be inferred (query variables), while factor nodes define the relationships between them based on local functions. Factor graphs (FGs) are graphical models that express the relationships between variables using two types of nodes: variable nodes and factor nodes. Each factor node corresponds to a function over a subset of variables and connects to multiple variable nodes to define their interdependencies. These factor functions are typically weighted to reflect their influence, with weights either manually set or learned from data. Due to their ability to model local dependencies effectively, FGs are widely applicable in solving optimization problems in robotics and artificial intelligence. The sum–product algorithm is commonly used in FGs to combine local functions into a global function, enabling inference across the network. This is particularly useful for computing optimal network capacity in tree-based topologies. In such structures, a global function encompasses the entire FG and supports efficient evaluation through message passing between nodes. The FG approach decomposes complex global functions into smaller, tractable components, represented in a bipartite graph format. Here, evidence variables have known values, while query variables represent unknowns to be inferred. Each factor node connects to variable nodes via factor functions, and its influence is quantified using a weight. These weights play a critical role in evaluating the overall network structure. Thanks to the locality property of many optimization challenges in robotics, FGs can model a wide array of AI problems. To identify the optimal root node that maximizes network capacity, the FG framework operates in three stages:

Shortest Path Spanning Trees (SPSTs) Construction: Each node in a wireless mesh network (WMN) is considered as a potential root, and Dijkstra's algorithm is applied to compute the shortest paths from that root to all other nodes. This step results in multiple SPST topologies—each rooted at a different node—with links assigned weights using the Airtime Link Metric (ALM). These SPSTs represent tree-structured topologies optimized for minimum transmission cost. ALM Link Weight Aggregation via Sum–Product Algorithm: For each SPST, the total link weight is calculated by summing the ALM values of individual links. While a simple summation might suffice, it fails to capture the network's complex interdependencies. Instead, the sum–product algorithm is used to compute the total ALM link weight more precisely by leveraging factor graph principles. It performs additions and multiplications across interconnected nodes, thereby considering the topology's structural influence. Within this process, all nodes are treated equally, and only the root node acts as a parent in the node hierarchy. The relationship among nodes is modeled multiplicatively. Optimal Root Node Selection: After computing the total ALM link weight for each SPST using the sum–product method, the root node associated with the lowest total cost is selected. This node forms the most efficient routing backbone for the network. This procedure is visually depicted in Step (3) of Figure 2, where the most optimal SPST topology is determined based on minimized link cost. By integrating factor graphs and the sum–product algorithm, this method offers a robust mechanism to evaluate and select optimal routing configurations in full-duplex wireless multihop networks, surpassing traditional metric aggregation techniques.

III. RESULTS AND DISCUSSION

This simulation-based evaluation investigates the performance of network capacity and computational time across two distinct scenarios. The first scenario assesses the effectiveness of the Factor Graph (FG) and Node Link Capacity (NLC) approaches, whereas the second focuses on analyzing the performance of the proposed Learning-based Path Selection (LPS) algorithms. The simulation assumes that each intermediate node is responsible for transmitting one packet to a central destination node. Nodes are designed to transmit and receive simultaneously, and can also handle two incoming packets at the same time.

The simulation environment was implemented in MATLAB R2022a, running on an Apple Mac mini (2018) with a 3.2 GHz Intel Core i7 processor and 64 GB of DDR4 RAM. In terms of the Q-learning setup, the initial learning rate was set to 1.0. However, due to modifications in the Q-function design specific to this study, a reevaluation of the optimal learning rate was necessary. As depicted in Figure 3, comparisons were conducted using network sizes of 50 and 100 nodes under consistent discount factors and threshold values. The analysis revealed that a learning rate of 1.0 prioritized only newly acquired information during Q-value updates, completely ignoring prior knowledge. To better simulate real-world conditions and incorporate historical data, the learning rate was adjusted to 0.9.

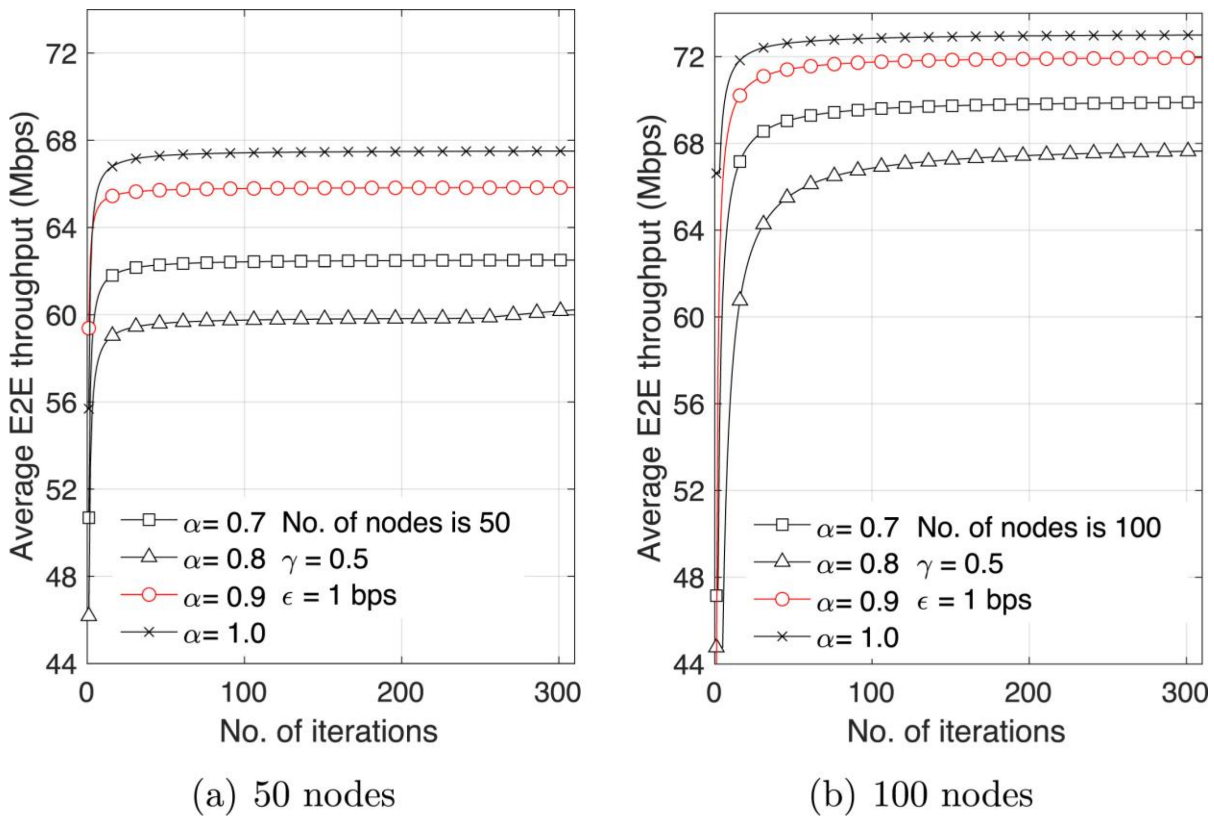
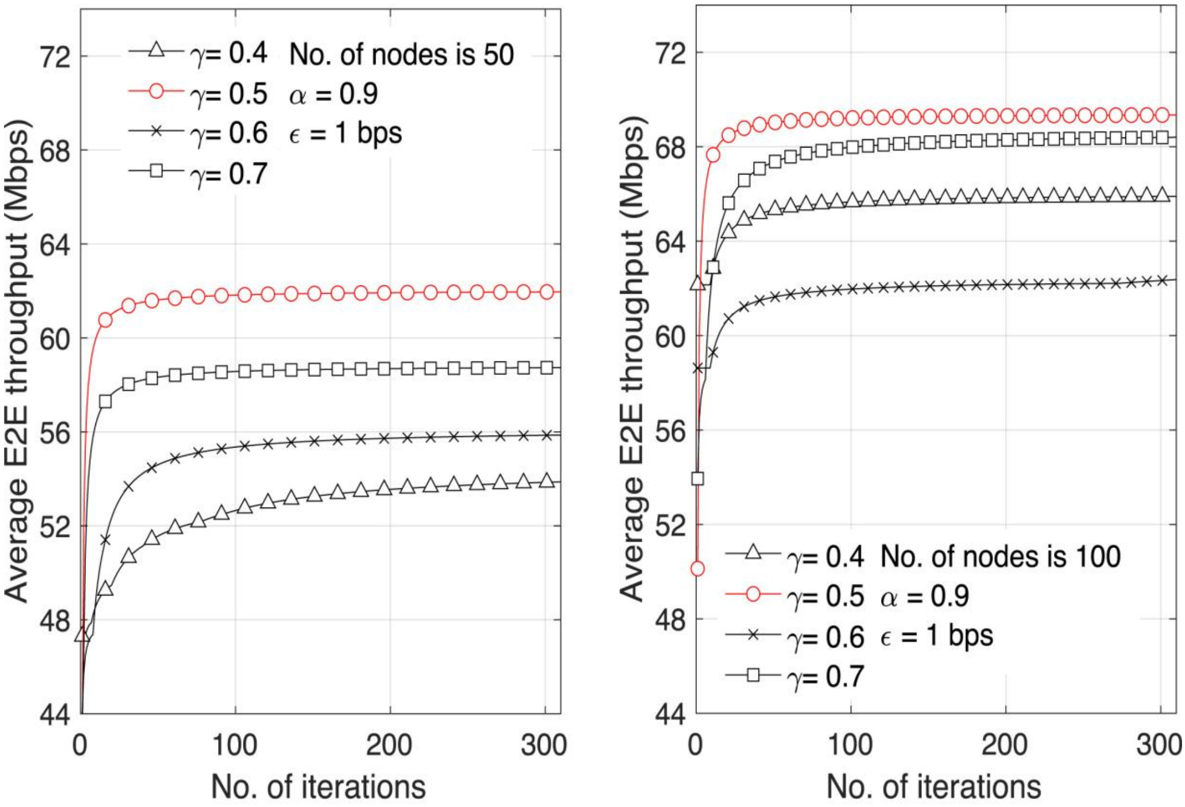
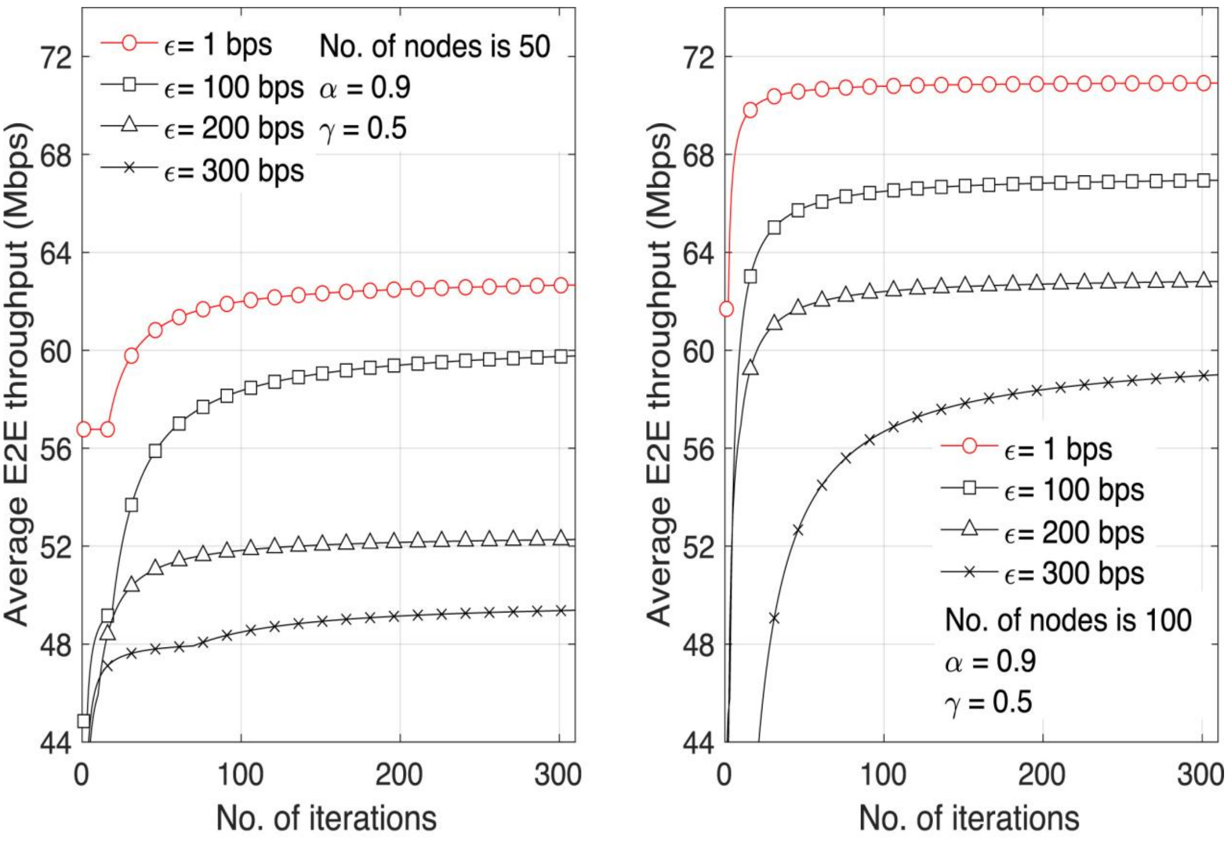


Fig. 3. Efficacy of learning rate in Q-learning



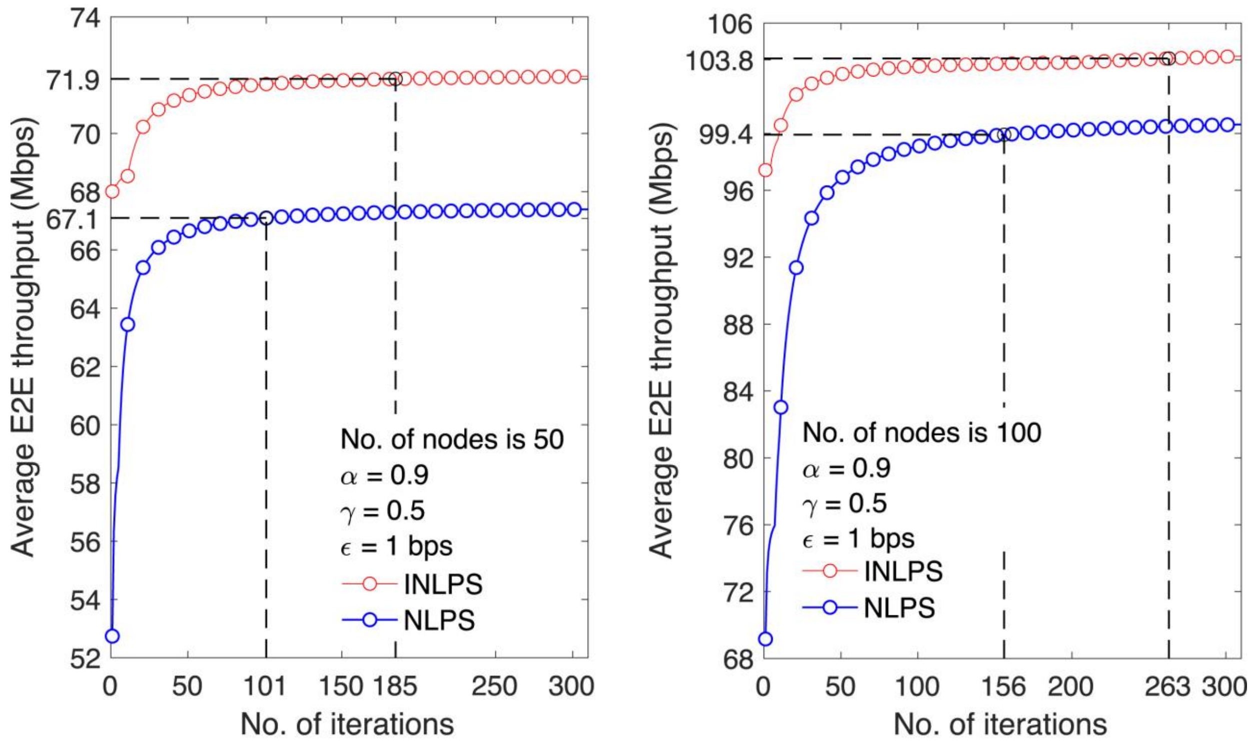
(a) 50 nodes (b) 100 nodes

Fig. 4. The discount factor's performance in Q-learning



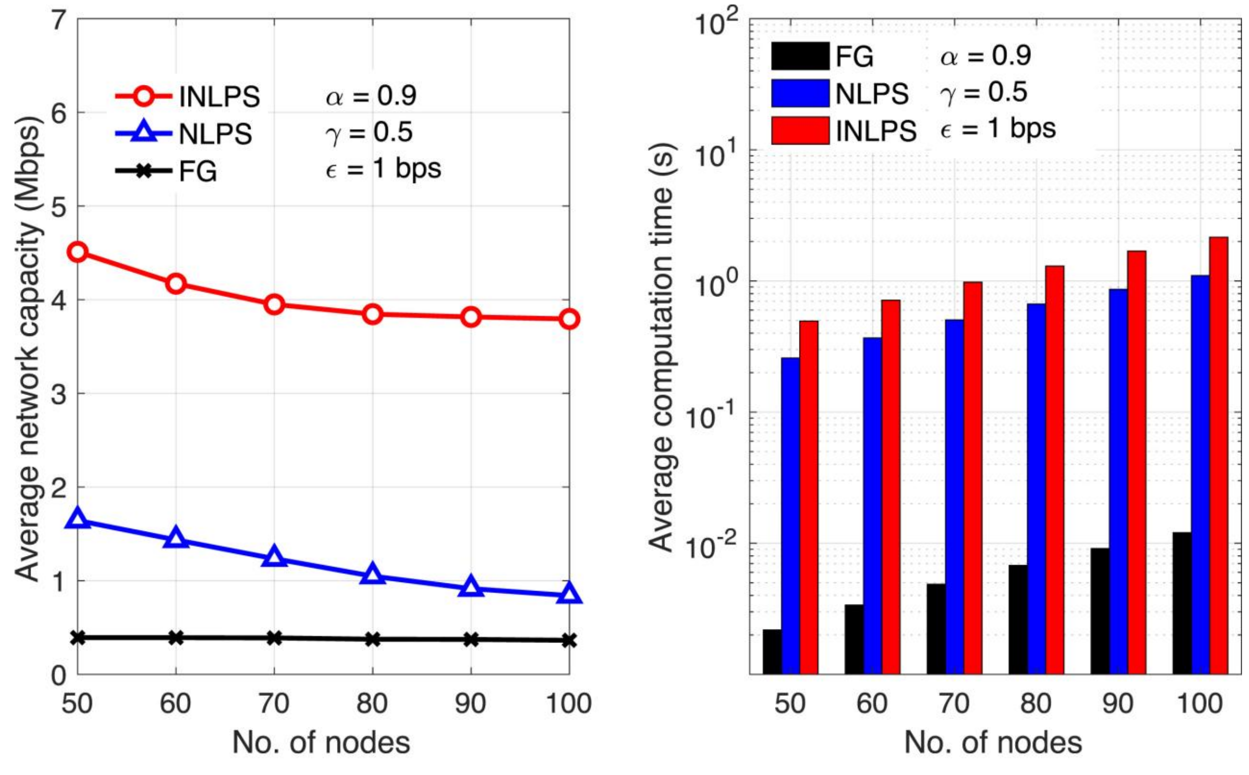
(a) 50 nodes (b) 100 nodes

Fig. 5. Q-learning threshold performance



(a) 50 nodes (b) 100 nodes

Fig. 6. Average E2E throughput against iterations for NLPS and INLPS algorithms



(a) Network capacity comparison (b) Computation time comparison

Fig. 7. Assessment of FG, NLPS, and INLPS performance

The discount factor in Q-learning reflects how much importance is assigned to future rewards when choosing actions. A higher discount factor implies that long-term gains are considered, whereas a lower factor favors short-term benefits. Figure 4 presents the results of varying discount factors, showing that a value of 0.5 maximizes average end-to-end throughput, making it the preferred setting for the training phase in this context. The Q-learning threshold determines when the training process concludes. Typically, smaller thresholds lead to more refined performance outcomes, although they may also extend the number of training iterations. As shown in Figure 5, testing with different threshold values in 50- and 100-node setups demonstrates that lower thresholds yield better average end-to-end throughput. When the threshold is set at 1 bps, training continues efficiently and stabilizes after approximately 300 iterations. Since throughput remains unchanged beyond that point, a threshold equivalent to 1 basis point is adopted for the final simulation configuration.

The numerical simulations are structured in two main phases to evaluate network capacity and computation time under various scenarios. In the first phase, we assess the performance of the system with and without the integration of Factor Graphs (FG) and Node Location Control (NLC), focusing on their influence on network capacity and computational efficiency. We also explore how employing FG aids in root node identification, contributing to the performance enhancement of the proposed factor graph-based Deep Reinforcement Learning (fDRL) model.

In the second phase, two algorithms—NLPS and its improved variant INLPS—are utilized to determine the optimal network topology within a Wireless Mesh Network (WMN). Specifically, the NLPS and INLPS configurations represent the use of their respective algorithms for path optimization, while the CoF (Combination of Factors) strategy is integrated with NLC to further enhance network throughput during data transmission.

Figure 6 highlights that the INLPS algorithm not only improves network performance but also manages power resources effectively. Although there is a slight reduction in average end-to-end (E2E) energy consumption when compared to NLPS, the gain in network capacity does not drastically compromise energy efficiency. This observation underscores the importance of simultaneously addressing energy conservation and performance optimization in network protocol design.

The results of network capacity analysis, shown in Figure 7(a), demonstrate that as the number of nodes increases from 50 to 100, overall capacity tends to decline. For instance, at 50 nodes, NLPS achieves an average throughput of 1.64 Mbps, while FG alone reaches only 0.39 Mbps. In contrast, INLPS achieves 4.51 Mbps, outperforming NLPS by a factor of 2.75 and FG by approximately 11.56 times. At 100 nodes, the capacities recorded are 3.79 Mbps (INLPS), 0.84 Mbps (NLPS), and 0.36 Mbps (FG). This substantial improvement is attributed to the INLPS algorithm's ability to refine Q-values using prior results from NLPS and apply Signal-to-Interference-plus-Noise Ratio (SINR) as a reward metric, allowing it to avoid low-SINR routes. Figure 7(b) presents the comparison of computation times, revealing that INLPS takes nearly twice the time required by NLPS when using Q-values as input, reflecting its more comprehensive optimization process. Since both algorithms incorporate FG during processing, FG efficiently determines the optimal root node, reducing the overhead. Without FG, each node would be compelled to independently execute NLPS or INLPS, making network topology optimization computationally expensive.The integration of FG significantly aids the NLPS and INLPS algorithms in identifying near-optimal network configurations, thereby achieving higher capacity with reduced processing time.

TABLE 1
COMPARATIVE ANALYSIS OF NETWORK CAPACITY AND COMPUTATION TIME FOR DIFFERENT ALGORITHMS

Algorithm	Avg. Network Capacity (Mbps)	Avg. Computation Time (s)	Throughput Improvement (vs. FG)	Time Efficiency (vs. FG)
FG	0.82	0.006	1×	1×
NLPS	3.52	0.627	4.29×	0.009×
INLPS	4.26	1.221	5.20×	0.004×

Table 1 highlights the performance comparison of three algorithms (FG, NLPS, INLPS) in terms of average network capacity and computation time based on simulation over a network of 100 nodes. The throughput improvement and time efficiency are calculated relative to the FG method. While FG shows superior computation time, both NLPS and INLPS significantly outperform in terms of network capacity.

IV. CONCLUSION

To enhance network capacity and minimize overall computational overhead, this study introduces a factor graph-based deep reinforcement learning (fDRL) framework for wireless mesh networks (WMNs). The proposed fDRL approach incorporates two Q-learning-based link path selection (LPS) algorithms: Named Link Path Selection (NLPS) and Improved Named Link Path Selection (INLPS). These algorithms utilize a root node, optimally selected through a factor graph (FG), to determine the most efficient path for each source node to reach the root. Leveraging FG significantly reduces training time, thereby improving the overall capacity performance of the network. Additionally, the fDRL framework employs a Combination of Functions (CoF) strategy to enable message transmission using fewer time slots and utilizes Network Linear Coding (NLC) for encoding and decoding, further boosting network efficiency. Compared to the traditional Tree-Based Routing (TBR) method, the FG-driven model can reduce computation time by up to 99%. When the network comprises 100 nodes, INLPS achieves a 4.26-fold improvement in capacity over NLPS, and a 5.08-fold increase relative to FG alone. Simulation results demonstrate that with 156 and 263 iterations for NLPS and INLPS respectively, the system achieves approximately 98% throughput while maintaining low computation time. These outcomes suggest that the proposed fDRL system is well-suited for integration with 6G wireless LAN extended service sets, particularly in scenarios involving multiple indoor devices forming a tree topology to transmit data to a central root node.

REFERENCES

[1] Alsharif M.H., Kelechi A.H., Albreem M.A., Chaudhry S.A., Zia M.S., Kim S. Sixth generation (6G) wireless networks: Vision, research activities, challenges and potential solutions Symmetry, 12 (4) (2020), pp. 1-5, 10.3390/sym12040676

[2] Trivedi G., Jabbari B. On wireless link connectivity for resilient multi-hop networks IEEE Mil. Commun. Conf., MILCOM (2021), pp. 285-290, 10.1109/MILCOM52596.2021.9652882

[3] Fujimura C., Sanada K., Mori K. Throughput analysis for string-topology full-duplex multi-hop network Int. Symp. on Wireless Pers. Multimedia Commun., WPMC (2017), pp. 535-541, 10.1109/WPMC.2017.8301871
[4] Rezaei S., Gharib M., Movaghar A. Throughput analysis of IEEE 802.11 multi-hop wireless networks with routing consideration: A general framework IEEE Trans. Commun., 66 (11) (2018), pp. 5430-5443, 10.1109/TCOMM.2018.2848905
[5] Lee W.J., Kim J.K., Kim K.R., Kim J.H. Second receiver selection algorithm for fairness in full duplex communications IEEE VTS Asia Pacific Wireless Commun. Symp., APWCS (2019), pp. 1-5, 10.1109/VTS-APWCS.2019.8851639
[6] Gui J., Zhou K. Flexible adjustments between energy and capacity for topology control in heterogeneous wireless multi-hop networks J. Netw. Syst. Manage., 24 (4) (2016), pp. 789-812, 10.1007/s10922-016-9367-y
[7] Baddu Naik Bhukya, V. Venkataiah, S. Mani.Kuchibhatla, S. Koteswari, R V S Lakshmi Kumari, and Yallapragada Ravi Raju, "Integrating the Internet of Things to Protect Electric Vehicle Control Systems from Cyber Attacks," IAENG International Journal of Applied Mathematics, vol. 54, no. 3, pp433-440, 2024
[8] Yu Y., Shah S., Tan Y., Lim Y. End-to-end throughput evaluation of consensus TPC algorithm in multihop wireless networks Int. Wireless Commun. and Mobile Comput. Conf., IWCMC (2015), pp. 941-946, 10.1109/IWCMC.2015.7289209
[9] Khun A.T.P., Lim Y., Tan Y. Optimal achievable transmission capacity scheme with transmit power control in full-duplex wireless multi-hop networks Int. Conf. on Mobile Comput. and Ubiquitous Netw., ICMU (2021), pp. 1-6, 10.23919/ICMU50196.2021.9638817
[10] Khun A.T.P., Lim Y., Tan Y. MAC protocol design and analysis for full-duplex wireless networks using MCST scheme Int. Wireless Commun. and Mobile Comput., IWCMC (2022), pp. 1394-1399, 10.1109/IWCMC55113.2022.9824398
[11] Mao Y., Kschischang F., Li B., Pasupathy S. A factor graph approach to link loss monitoring in wireless sensor networks IEEE J. on Sel. Areas in Commun., 23 (4) (2005), pp. 820-829, 10.1109/JSAC.2005.843557
[12] Li W., Yang Z., Hu H. Distributed multi-sensor tracking in wireless networks using nonparametric variant of sum-product algorithm Asia-Pacific Conf.on Commun., APCC (2013), pp. 132-137, 10.1109/APCC.2013.6765929
[13] Baddu Naik Bhukya, Padmanabha Raju Chinda, Srinivasa Rao Rayapudi, and Swarupa Rani Bondalapati, "Advanced Control with an Innovative Optimization Algorithm for Congestion Management in Power Transmission Networks," Engineering Letters, vol. 31, no.1, pp194-205, 2023
[14] Bu X., Liu C., Yu Q., Yin L., Tian F. Optimization on cooperative communications based on network coding in multi-hop wireless networks Int. Wireless Commun. and Mobile Comput., IWCMC (2020), pp. 384-387, 10.1109/IWCMC48107.2020.9148060
[15] Xue J., Kurkoski B.M. Lower bound on the error rate of genie-aided lattice decoding IEEE Int. Symp. on Inf. Theory, ISIT (2022), pp. 3232-3237, 10.1109/ISIT50566.2022.9834527
[16] Rosenberger J., Urlaub M., Rauterberg F., Lutz T., Selig A., Bühren M., Schramm D. Deep reinforcement learning multi-agent system for resource allocation in industrial Internet of Things Sensors, 22 (11) (2022), p. 4099, 10.3390/s22114099
[17] Wang Y., Shang F., Lei J. Reliability optimization for channel resource allocation in multihop wireless network: A multigranularity deep reinforcement learning approach IEEE Internet of Things J., 9 (20) (2022), pp. 19971-19987
[18] Ho T., Nguyen K. Joint server selection, cooperative offloading and handover in multi-access edge computing wireless network: A deep reinforcement learning approach IEEE Trans. Mobile Comput., 21 (7) (2022), pp. 2421-2435, 10.1109/TMC.2020.3043736
[19] Mazhar T., Haq I., Ditta A., Mohsan S.A.H., Rehman F., Zafar I., Gansau J.A., Goh L.P.W. The role of machine learning and deep learning approaches for the detection of skin cancer Healthcare, 11 (3) (2023), p. 415
[20] Lee Y., Ma X., Lang S.A., Valderrama-Araya F.E., Chapuis L.A. Deep learning based modeling of wireless communication channel with fading (2023) arXiv preprint arXiv:2312.06849
[21] Dugaev D.A., Matveev I.G., Siemens E., Shuvalov V.P. Adaptive reinforcement learning-based routing protocol for wireless multihop networks Int. Scientific-Tech. Conf. on Actual Problems of Electron. Instrum. Eng., APEIE (2018), pp. 209-218, 10.1109/APEIE.2018.8545412
[22] Wongphatcharatham T., Phakphisut W., Wijitpornchai T., Areepray-oonkij P., Jaruvitayakovit T., Hannanta-anan P. Multi-agent Q-learning for power allocation in interference channel Int. Tech. Conf. on Circuits/Syst., Comput. and Commun., ITC-CSCC (2022), pp. 876-879, 10.1109/ITC-CSCC55581.2022.9894852
[23] Wang X., Wang X. Reinforcement learning-based multihop relaying: A decentralized Q-learning approach Entropy (Basel), 23 (10) (2021), 10.3390/e23101310
[24] Su Y., Lu X., Zhao Y., Huang L., Du X. Cooperative communications with relay selection based on deep reinforcement learning in wireless sensor networks IEEE Sens. J., 19 (20) (2019), pp. 9561-9569, 10.1109/JSEN.2019.2925719 View at publisherView in
[25] ANSI/IEEE Standard 802.11 Y. Part II: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications mesh networking IEEE Standard Specification 802.11s Mesh WLAN (2011), pp. 1-372, 10.1109/IEEESTD.2011.6018236.