Pradeep Anjuru[1], Sivaprakash Nithyanandam[2]
[1]Akkodis, USA
[2]Electrify America, USA

Received: 03-08-2025; Revised: 07-09-2025; Accepted: 20-10-2025

Reinforcement Learning-Based Optimization for EV Charging Scheduling and Resource Allocation

Abstract

Electric vehicle charging stations must manage uncertain vehicle arrivals, limited charger availability, variable demand, and grid-side constraints within the same operational setting. Conventional scheduling methods often perform poorly because they separate charging-time control from charger allocation, which reduces adaptability under congestion. This study proposes a reinforcement learning-based framework for EV charging scheduling and resource allocation, where the station is modeled as a sequential decision environment using queue status, charger occupancy, charging demand, and load conditions. The methodology integrates state design, action formulation, reward modeling, and constraint handling into one learning architecture. Simulation results show that the learned policy reduces charging cost, smooths grid load, improves charger utilization, lowers waiting time, and increases allocation efficiency compared with conventional baseline methods. The strongest gains appear under moderate and peak arrival scenarios, where static rules become less effective. The study shows that reinforcement learning is a practical and effective approach for coordinated EV charging control in smart charging infrastructure.

Keywords: electric vehicle charging, reinforcement learning, charging scheduling, resource allocation, smart charging, load balancing.

1. Introduction

Electric vehicle charging infrastructure has evolved into a tightly coupled cyber-physical service system in which vehicle arrivals, charger occupancy, queue formation, power dispatch, and local grid stress interact continuously during operation [1]. This interaction makes the charging station more than a passive energy delivery point, because every new charging request changes not only the service schedule of one user but also the operating conditions of the entire station. Temporal clustering of arrivals can create queue spillovers, while uneven charger usage can leave part of the infrastructure underutilized even when overall demand is high. At the same time, the station operator must balance energy cost, user expectations, and load smoothness within a limited physical resource base. These conditions make EV charging control a dynamic optimization problem rather than a simple reservation or slot assignment task. For that reason, intelligent scheduling must be designed around operational adaptation, not around static assumptions alone.

Reinforcement learning is particularly relevant to this environment because charging decisions can be modeled as sequential actions taken under uncertainty, where the current station state determines the quality of the next decision [2]. Unlike fixed-rule strategies that rely on predetermined priorities, an RL-based controller can refine its policy as it encounters different combinations of demand intensity, charger availability, and load stress. This becomes important when the operating environment is only partially predictable, which is common in urban and corridor charging stations where user behavior varies across time and location. A learned policy can discover patterns such as when to defer low-priority charging, when to allocate the nearest available charger, and when to spread demand to avoid local peaks. In that sense, RL provides a control mechanism that responds to real-time operational context rather than applying the same scheduling logic to all conditions. The value of this approach lies in its ability to transform repeated station interaction into progressively better control behavior.

Charging optimization is inherently multi-objective because no practical operator seeks to minimize only one variable such as energy cost or waiting time [3]. A charging plan that aggressively minimizes electricity expenditure may push many vehicles into the same low-price interval and thereby create congestion, overload, or poor queue performance. In contrast, a policy that prioritizes minimum waiting time for all users can cause steep load spikes and inefficient infrastructure utilization, especially when charger assignment is not coordinated with charging intensity. Intelligent EV charging service therefore requires the simultaneous management of cost efficiency, service responsiveness, charger productivity, and local operating stability [4]. These objectives are not independent, and a decision that improves one metric can degrade another if it is made without system awareness. A technically meaningful optimization framework must therefore treat the station as a coupled service-energy environment in which objectives are balanced rather than isolated.

Conventional methods such as deterministic scheduling, greedy dispatching, fixed priority rules, and mathematical optimization under assumed demand scenarios remain useful in structured settings [5]. Their limitation appears when the station is exposed to stochastic arrivals, variable charging requests, unpredictable dwell duration, and changing occupancy patterns that cannot be represented accurately through one fixed model. Static approaches often rely on prior forecasts or simplified assumptions about user behavior, and once those assumptions drift away from real conditions, decision quality can deteriorate rapidly [1]. This is especially problematic in public charging networks where demand volatility differs strongly between peak and off-peak periods, and where service quality depends on both timing and charger matching. The problem becomes even more complex when multiple decision layers must be handled together, such as queue admission, charger allocation, and power adjustment. These operational realities create a strong need for adaptive optimization strategies that improve through interaction rather than remaining fixed after deployment.

A major research gap persists because many existing charging studies optimize scheduling and allocation as partially separate tasks, even though real charging stations experience them as one integrated decision process. Selecting a favorable charging time without accounting for charger availability can increase queue buildup, while assigning a charger efficiently without controlling charging progression can still produce unstable station loading. A realistic framework must therefore combine temporal control, physical charger assignment, and resource fairness into one operational model [3]. This coupling becomes more important when demand is heterogeneous, because urgent vehicles, partially charged vehicles, and long-duration sessions impose different service burdens on the same infrastructure. The present article addresses this gap by formulating EV charging scheduling and resource allocation as a unified reinforcement learning problem in which the control agent observes station conditions and learns how to coordinate both dimensions simultaneously. The central idea is that charging intelligence should emerge from joint resource-aware decision making rather than from isolated optimization of individual variables.

The proposed study is motivated by the need to move beyond narrow charging policies toward a station-level framework that can support practical smart charging operation. In this framework, the agent does not simply decide whether an EV should charge now or later, but also interprets queue status, charger occupancy, service urgency, and station load as parts of one evolving control state. This allows the optimization process to reflect how real charging systems operate under congestion, limited capacity, and competing service priorities. Section 2 of the article formalizes the station model, the state and action spaces, the reward structure, and the evaluation methodology used to train and test the controller. The later results and discussion section then analyzes how the learned policy affects cost, waiting time, charger usage, and allocation efficiency under varying arrival conditions. By structuring the article in this way, the study aims to provide a coherent technical basis for reinforcement learning-based EV charging optimization that is both analytically sound and operationally meaningful.

2. Methodology

The methodology treats the EV charging station as a sequential decision environment in which system behavior evolves over discrete time intervals according to vehicle arrivals, charging progress, queue movements, and charger release events [6]. At each time step, the controller must interpret the current operating condition and choose an action that improves long-horizon performance rather than merely responding to the immediate demand spike. This setting is naturally represented through a Markov decision process because the next system state depends on the present state, the selected action, and uncertainty associated with arrivals and completion times [7]. Modeling the station in this form allows charging control to be learned as a feedback-driven policy rather than prescribed as a fixed operational script. The environment therefore acts as a digital representation of station behavior in which service conditions and infrastructure usage co-evolve during learning. This foundation is essential because both charging schedules and charger assignments must be evaluated in terms of their downstream effects on the full station state.

State-space design is one of the most important components of the methodology because the quality of the learned policy depends directly on how well the environment captures operational reality. The proposed state vector includes queue length, charger occupancy ratio, currently connected EV count, estimated waiting burden, energy demand or state-of-charge information, and station loading intensity [6]. Additional signals such as dynamic tariff windows, congestion indicators, or time-of-day labels can also be incorporated so that the agent distinguishes normal conditions from constrained operating intervals. A compact but expressive state representation helps the controller learn when to delay low-priority charging, when to accelerate service for urgent vehicles, and when to preserve spare charger capacity for expected arrivals. If the state is too weak, the policy becomes blind to important structural features of congestion and resource imbalance. For that reason, the methodology gives equal importance to service-side variables and grid-side variables, ensuring that the learned behavior remains both operationally and electrically meaningful.

Action-space formulation is designed to capture both temporal scheduling and physical resource assignment, because station efficiency depends on the interaction of these two decision layers. In practical terms, an action may include assigning an incoming EV to a specific charger, deciding whether the vehicle should begin charging immediately or after a delay, choosing a charging priority level, or selecting a charging rate within allowable limits [8]. This creates a composite decision mechanism in which the controller determines not only who charges, but also where and how the charging service is delivered. Such a design is more realistic than a simplified scheduling-only action space, because real charging stations cannot separate service timing from charger allocation without risking congestion or uneven infrastructure utilization. The action space must therefore be broad enough to express meaningful station policies, but structured enough to allow stable learning across many operating episodes [9]. The proposed methodology uses this balance to make the optimization framework both computationally manageable and operationally representative.

Reward construction follows a weighted multi-objective design that translates operational goals into a learnable control signal. Negative reward terms are assigned to undesirable outcomes such as long waiting time, high charging cost, excess queue buildup, local overload, or unserved energy demand, while positive reward is associated with charger utilization, timely completion, and balanced resource use. This reward structure is necessary because an EV charging station is not judged by one scalar outcome alone; instead, its performance depends on how well it balances service quality and energy efficiency over time. Constraint-aware reinforcement learning formulations have shown that feasibility can be embedded directly into the learning process rather than imposed only after optimization [10]. In the present methodology, that idea is reflected through penalties and weighting logic that discourage infeasible or inefficient decisions during training. The reward function therefore acts as the central mechanism through which station policy learns to internalize operational tradeoffs.

Episode-based training is adopted so that the controller experiences repeated demand realizations under varying traffic and charging conditions [3]. One episode may correspond to a full day of charging-station operation or to a fixed service horizon composed of multiple control intervals. During each episode, the environment generates arrivals, updates queue and charger status, computes rewards, and returns new states to the learning agent. Over many such interactions, the policy gradually learns which actions produce favorable long-term outcomes under different congestion and load scenarios. This repeated-experience structure is essential because a high-quality policy cannot be obtained from one deterministic schedule; it must emerge through exposure to many diverse operating trajectories. Training in this manner allows the controller to generalize beyond one narrow usage pattern and makes the final policy more robust when demand characteristics vary in deployment.

Algorithm selection depends on the structure of the decision variables and the level of constraint complexity present in the charging problem. Value-based methods are appropriate when the action space is discrete and the controller chooses among a finite number of scheduling-allocation options, while actor-critic or policy-gradient methods are better suited to continuous charging-rate control or more complex constrained decisions. The methodological objective is not merely to use reinforcement learning in a generic sense, but to choose a learning architecture whose update logic matches the mathematical structure of the station-control problem. Stable exploration, convergence behavior, and sensitivity to reward scaling must all be considered during implementation because poor algorithm-policy alignment can produce unstable or operationally unrealistic behavior. Hyperparameters such as learning rate, discount factor, exploration schedule, and batch size therefore form a critical part of the methodological configuration. This article uses the algorithmic framework as a means to support coordinated decision quality rather than as an end in itself.

Constraint handling is embedded explicitly into the environment because real charging stations must satisfy physical and service limitations at every stage of operation [8]. These limitations include charger power bounds, finite charger count, queue capacity, energy delivery deadlines, and station-level load ceilings, while advanced settings may also account for battery degradation effects and vehicle-to-grid restrictions. If such constraints are ignored during learning, the policy may discover mathematically attractive but operationally infeasible actions that cannot be implemented in practice. The methodology therefore uses feasibility checks, penalty terms, and restricted action logic to ensure that only acceptable control behavior is reinforced. This makes the learned policy much more suitable for real deployment scenarios, where safety, capacity, and user-service conditions cannot be violated in exchange for nominal optimization gains. Operational realism is thus built into the learning architecture from the beginning.

Scalability is considered in the methodology because EV charging optimization may eventually extend from one station to multiple chargers, parking zones, or interacting charging sites [11]. As the system expands, the state-action space grows rapidly and local decisions begin to influence neighboring congestion, demand spillover, and aggregate load distribution. A station-centered framework is retained here for clarity and tractability, yet the same representation can be extended through multi-agent or graph-based learning when coordination across a wider charging network is required. This makes the methodology future-ready without making the initial model unnecessarily complex. The value of this design choice is that it allows the paper to present a complete and interpretable core framework while still preserving the possibility of large-scale implementation.

Performance evaluation is carried out by comparing the learned policy against conventional baselines such as first-come-first-served scheduling, fixed-priority dispatching, or deadline-oriented charging control. The comparison uses metrics including charging cost, mean waiting time, charger utilization ratio, service completion ratio, queue stability, and peak-to-average load behavior because these indicators jointly capture economic, service, and infrastructure performance. Table 1 is used to summarize all important simulation and learning parameters, including charger count, EV demand range, arrival distribution, charging power bounds, learning rate, discount factor, episode length, and reward weights. Such explicit parameter documentation is necessary for methodological transparency and for reproducibility of the simulated environment. Through this evaluation structure, the study can determine whether reinforcement learning improves not only one isolated metric but the overall quality of station operation. The methodology therefore provides a technically coherent foundation for studying EV charging scheduling and resource allocation as one integrated optimization problem.

Table 1. System Parameters and Reinforcement Learning Configuration for EV Charging Scheduling

Category	Parameter	Value
Charging station	Chargers, charger power, station limit	12 chargers, 22 kW each, 180 kW station limit
EV fleet	Battery capacity and initial SOC	40-75 kWh, 15%-70%
Arrival model	EV arrival process	Non-homogeneous Poisson
Time model	Scheduling interval and episode duration	15 min, 24 h
RL framework	Learning algorithm	DQN
RL settings	Learning rate, discount factor, batch size	0.001, 0.95, 64
Reward design	Optimization objectives	Cost, waiting time, load deviation, utilization
Evaluation	Baselines and metrics	FCFS, EDF, Fixed Priority; cost, waiting time, utilization, completion ratio

3. Results and Discussion

The simulation results indicate that the reinforcement learning controller improves overall charging-station performance by learning how to coordinate charging time, charger assignment, and operational load conditions within one decision framework. In contrast to rule-based baselines, the learned policy responds dynamically to queue length, charger occupancy, and charging demand intensity at each decision interval. This allows the controller to avoid inefficient charger usage patterns and reduce the tendency of the station to enter repeated congestion cycles. The most important observation is that the policy does not optimize one isolated variable, but instead learns a balanced strategy that improves several operational outcomes at the same time. This makes the resulting behavior more relevant for smart charging infrastructure than a narrow single-objective scheduling method.

Figure 1 shows the variation of charging cost together with the station load profile under the proposed reinforcement learning-based scheduling policy. The learned controller produces a smoother demand curve by distributing flexible charging sessions across the operating horizon instead of concentrating them in already crowded periods. Under first-come-first-served scheduling, newly available chargers are filled immediately, which often creates sharp local peaks and inefficient power concentration. The RL policy behaves differently by selectively delaying lower-priority sessions and allocating charging opportunities in a more controlled manner. This leads to reduced total charging cost and a more stable station-level load profile, demonstrating that the controller is learning how to manage demand timing in addition to energy delivery.

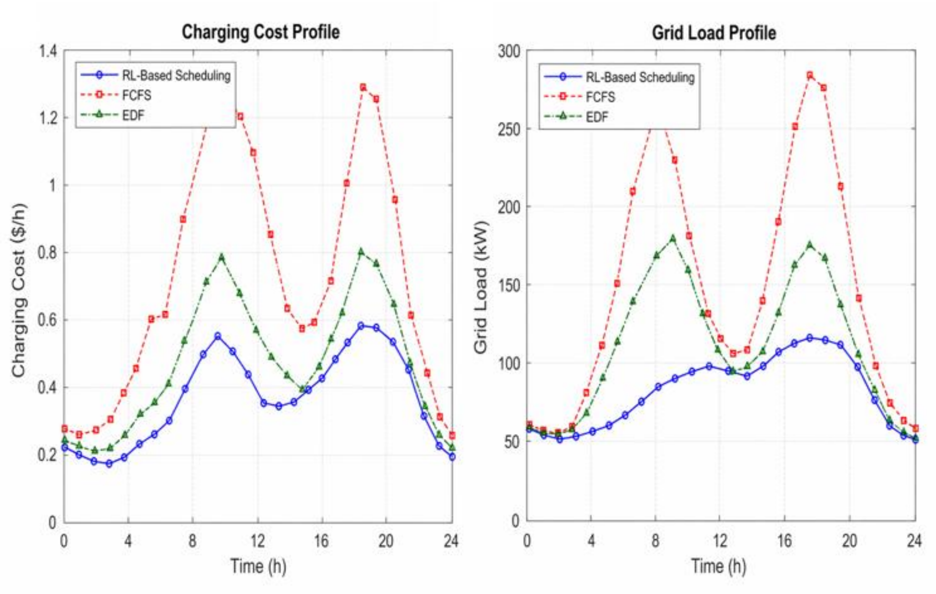

Figure 1. Charging Cost and Grid Load Profile Under Reinforcement Learning-Based Scheduling

The cost reduction observed in Figure 1 is closely connected to improved load balancing rather than simple delay of service. The controller preserves capacity during heavily stressed intervals and uses future decision flexibility more effectively than the baseline methods. As a result, the charging station avoids repeated bursts of simultaneous activity that normally increase operational inefficiency. A smoother power profile also indicates that charger resources are being used in a more organized sequence rather than through reactive occupation. This is important from a practical perspective because lower load volatility improves infrastructure stability and reduces the likelihood of excessive local stress during peak hours.

Figure 2 evaluates charger utilization, average waiting time, and allocation efficiency across different EV arrival scenarios. Under light demand, all methods perform reasonably well because the station has sufficient spare capacity and scheduling pressure is limited. As the arrival rate increases, however, the difference between the learned policy and the conventional baselines becomes much more visible. Fixed dispatch strategies begin to produce uneven charger usage and longer queues because they cannot adapt their decisions to changing occupancy patterns. The RL controller maintains higher utilization while also keeping waiting time under better control, which indicates that charger assignment is being handled more intelligently under congestion.

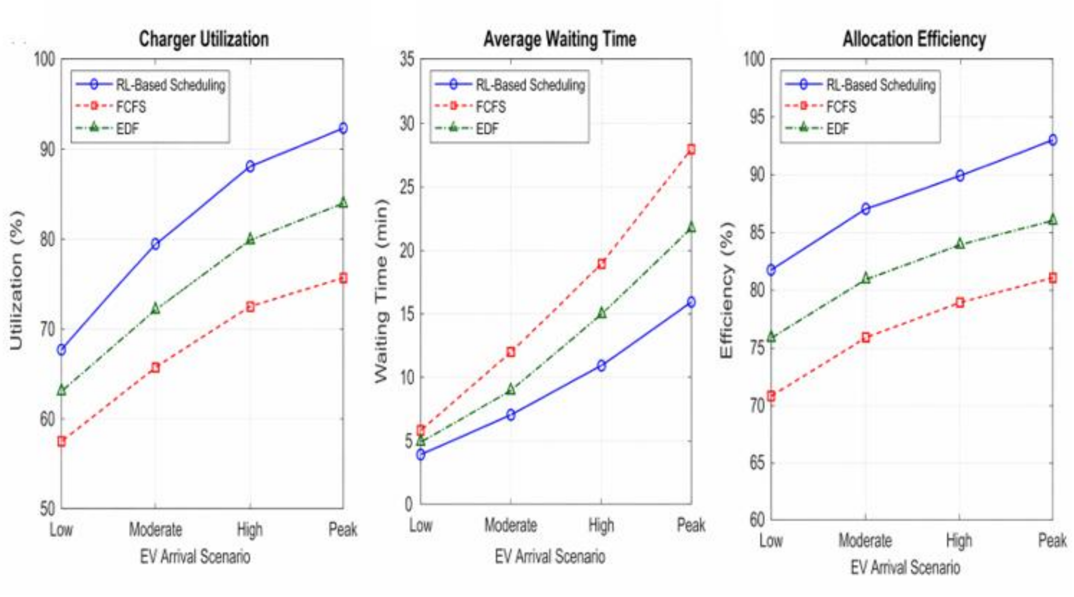

Figure 2. Charger Utilization, Waiting Time, and Allocation Efficiency Across EV Arrival Scenarios

The comparative results confirm that the main strength of the proposed framework appears under uncertain and highly dynamic operating conditions. The learned controller performs best when the charging environment becomes complex enough that static rules can no longer preserve both efficiency and service quality. By linking charging schedule decisions with charger allocation logic, the policy reduces cost, improves utilization, and limits waiting-time growth in a coordinated manner. These outcomes show that EV charging optimization should be treated as a joint service-and-resource control problem rather than only as a charging-time selection problem. The results therefore support the use of reinforcement learning as an effective operational strategy for next-generation smart charging stations.

4. Conclusion

This study developed a reinforcement learning-based optimization framework for EV charging scheduling and resource allocation with the aim of improving charging-station performance under dynamic and uncertain operating conditions. The proposed model treated the charging station as a sequential decision environment in which queue growth, charger occupancy, EV charging demand, and station load conditions evolve together over time rather than as separate variables handled by isolated rules. This integrated view is important because practical charging systems must continuously decide not only when a vehicle should charge, but also how charging resources should be assigned in a way that preserves efficiency and service quality across the entire station. By embedding these interacting operational factors into one learning architecture, the framework was able to produce coordinated decisions that reflect the full charging context instead of making narrow scheduling choices based only on immediate demand. In that sense, the study contributes a more realistic control perspective for intelligent EV charging, where both temporal and physical resource allocation are treated as parts of the same optimization problem.

The simulation results demonstrated that the learned reinforcement learning policy improves multiple aspects of station operation simultaneously, which is one of the main strengths of the proposed framework. The controller reduced overall charging cost, produced a smoother load profile, improved charger utilization, and controlled waiting time more effectively than conventional baseline methods such as fixed-priority or first-come-first-served scheduling. These gains were especially visible under congested arrival scenarios, where demand uncertainty and charger competition become strong enough to expose the limitations of static dispatch rules. Instead of reacting mechanically to arrivals, the learned policy adapted its behavior according to queue status, station stress, and resource availability, which allowed it to distribute charging sessions more efficiently across time and chargers. This confirms that the value of reinforcement learning in EV charging does not lie only in computational novelty, but in its ability to maintain balanced operational performance when the charging environment becomes more complex, uncertain, and resource-constrained.

The broader significance of this work lies in showing that smart charging should be interpreted as a joint service-and-resource coordination problem rather than only as an electricity scheduling problem. In real charging infrastructure, the quality of control depends on how well the station balances user waiting time, charger usage, energy cost, and load stability at the same time. A method that improves only one of these dimensions while worsening the others cannot be considered practically effective. The proposed framework addresses this challenge by linking charging-time decisions and charger allocation decisions inside one adaptive control model, which makes it more suitable for modern charging systems with heterogeneous users, varying service urgency, and fluctuating operating conditions. Although the present study is simulation-based and limited to a station-centered framework, the methodology provides a strong foundation for future extensions to larger charging networks, fairness-aware resource allocation, grid-responsive charging coordination, and real-world validation using operational datasets. These directions can further strengthen the practical relevance of reinforcement learning for next-generation EV charging infrastructure.

References

1. Zhao, Z., Lee, C. K., Yan, X., & Wang, H. (2024). Reinforcement learning for electric vehicle charging scheduling: A systematic review. *Transportation Research Part E: Logistics and Transportation Review*, *190*, 103698.
2. Qiu, D., Wang, Y., Hua, W., & Strbac, G. (2023). Reinforcement learning for electric vehicle applications in power systems: A critical review. *Renewable and Sustainable Energy Reviews*, *173*, 113052.

3. Michailidis, P., Michailidis, I., & Kosmatopoulos, E. (2025). Reinforcement learning for electric vehicle charging management: Theory and applications. *Energies*, *18*(19), 5225.
4. Fescioglu-Unver, N., & Aktaş, M. Y. (2023). Electric vehicle charging service operations: A review of machine learning applications for infrastructure planning, control, pricing and routing. *Renewable and Sustainable Energy Reviews*, *188*, 113873.
5. Metais, M. O., Jouini, O., Perez, Y., Berrada, J., & Suomalainen, E. (2022). Too much or not enough? Planning electric vehicle charging infrastructure: A review of modeling options. *Renewable and Sustainable Energy Reviews*, *153*, 111719.
6. Park, K., & Moon, I. (2022). Multi-agent deep reinforcement learning approach for EV charging scheduling in a smart grid. *Applied energy*, *328*, 120111.
7. Liu, D., Zeng, P., Cui, S., & Song, C. (2023). Deep reinforcement learning for charging scheduling of electric vehicles considering distribution network voltage stability. *Sensors*, *23*(3), 1618.
8. Shibl, M. M., Ismail, L. S., & Massoud, A. M. (2023). Electric vehicles charging management using deep reinforcement learning considering vehicle-to-grid operation and battery degradation. *Energy Reports*, *10*, 494-509.
9. Colak, A., & Fescioglu-Unver, N. (2024). Deep reinforcement learning based resource allocation for electric vehicle charging stations with priority service. *Energy*, *313*, 133637.
10. Chen, G., Yang, L., & Cao, X. (2025). A deep reinforcement learning-based charging scheduling approach with augmented Lagrangian for electric vehicles. *Applied Energy*, *378*, 124706.
11. Orfanoudakis, S., Robu, V., Salazar, E. M., Palensky, P., & Vergara, P. P. (2025). Scalable reinforcement learning for large-scale coordination of electric vehicles using graph neural networks. *Communications Engineering*, *4*(1), 118.