Machine Learning-Based Predictive Medium Access Control Protocols for Vehicular Networks

IJACMET
International Journal of Advanced Computational Methodologies and Emerging Technologies
1–9

OwenPress

Nika Beridze[1] and Salome Kvitsiani[2]

Abstract

The design of robust Medium Access Control (MAC) protocols for vehicular networks remains a critical challenge due to dynamic topologies, intermittent channel conditions, and stringent latency requirements. This paper presents a machine learning-based predictive MAC framework that leverages historical and real-time channel state information to optimize contention window adaptation, priority scheduling, and collision avoidance. A hybrid architecture integrating recurrent neural networks (RNNs) with reinforcement learning (RL) agents is developed to predict temporal traffic patterns and dynamically adjust channel access parameters. The RNN module processes time-series data from vehicular nodes to forecast short-term network congestion, while the RL agent optimizes transmission policies through a discounted reward mechanism based on collision probability and throughput maximization. Extensive simulations under urban and highway scenarios demonstrate a 27% reduction in end-to-end latency and 33% improvement in packet delivery ratio compared to IEEE 802.11p and conventional CSMA/CA. However, the model exhibits a 15–18% performance degradation in non-line-of-sight (NLOS) environments with multipath fading, attributed to imperfect channel estimation. Further analysis reveals quadratic computational complexity relative to node density, limiting scalability beyond 150 nodes per coverage area. The proposed framework is shown to achieve a Nash equilibrium in transmission scheduling under stationary traffic, though convergence time increases exponentially with velocity variance. These results highlight the potential of cross-layer integration of machine learning into MAC protocols while underscoring the need for distributed computation to address real-time constraints.

Introduction

Vehicular ad-hoc networks face critical performance requirements due to the high mobility of vehicles, the diverse traffic patterns arising from multiple applications, and the stringent latency constraints inherent in safety-related messaging (*1*–*3*). Many of the existing medium access control protocols, particularly those based on the IEEE 802.11 family of standards, do not offer sufficient adaptability for the dynamic conditions of vehicular networks. One of the core difficulties lies in the variability of traffic load, where vehicles experience bursts of data transmission demands for infotainment, sensor sharing, and cooperative driving applications. Furthermore, higher mobility leads to rapidly changing network topologies, rendering purely reactive contention window adjustments insufficient and slow to adapt to ephemeral patterns. (*4*, *5*) The design of robust Medium Access Control (MAC) protocols for vehicular networks remains a critical challenge due to the highly dynamic nature of vehicular topologies, where nodes frequently join and leave the network, causing rapid topology changes that impact communication stability (*6*). Additionally, intermittent channel conditions, influenced by factors such as interference, fading, and obstacles in urban or rural environments, further complicate reliable data transmission. The premise of research into adaptive protocols is motivated by the failure of fixed backoff strategies to maintain optimal performance across a range of traffic intensities. With non-negligible probability, bursty sources that converge in the same region can provoke collisions and exponential backoff expansions. During extended contention resolution, critical safety messages can be delayed, which is obviously undesirable in applications such as collision avoidance or cooperative maneuvering (*7*).

[1]Akaki Tsereteli State University, Faculty of Informatics, Tamar Mepe Street 59, Kutaisi, Georgia
[2]Georgian Technical University, Department of Computer Systems and Networks, Kostava Street 77, Tbilisi, Georgia

Vehicle-to-everything communication systems must handle both high-rate sensor data from radar and LIDAR devices (8), as well as sporadic, bursty emergency beacons. An immediate reconfiguration in the backoff or contention window policy could avert collisions, but naive solutions suffer from partial or outdated observability of network conditions.

Recent advances in machine learning offer promising directions for addressing these limitations (9). In particular, reinforcement learning approaches that rely on a state-action-reward framework can incorporate a variety of contexts, such as recent collisions, inter-vehicle distances, and the history of received signal power or interference levels. Deep neural networks can combine heterogeneous inputs (like signal features, queue lengths, or velocity vectors) and learn a mapping to improved channel access decisions. Among these approaches, dueling Q-networks stand out for their ability to decouple the estimation of the value function from the advantage function, thereby stabilizing learning. However, there exist theoretical challenges in analyzing convergence when partial observability is introduced, especially under fast-changing vehicular conditions. (10)

Partial observability manifests in vehicular networks for two key reasons. First, each node possesses incomplete information about its neighbors' queues and decisions. Second, measurement noise and shadowing may obscure the exact channel state at different times and locations, leading to unreliable or outdated channel state information (11). Therefore, framing the MAC decision process as a partially observable Markov decision process (POMDP) has garnered attention, since it encapsulates the uncertainty in the state transitions and partial knowledge of the environment. Solving a POMDP optimally is generally computationally challenging, but approximate methods guided by neural networks can enable near-real-time decision making, particularly with well-designed state encodings and memory mechanisms.

One of the most promising families of neural architectures for capturing temporal dependencies is based on recurrent neural networks, such as LSTM (Long Short-Term Memory) (12). LSTMs can maintain hidden states for extended time intervals, enabling them to learn patterns in channel occupancy over time, especially in environments where bursts of interference or collisions exhibit correlation. For example, a sudden surge in transmissions from a platoon of vehicles might be predictable if the vehicles have just entered a congested highway segment or are engaged in cooperative braking. By foreseeing this surge, an adaptive MAC protocol can either reduce the contention window in anticipation of future idle periods or expand it preemptively to mitigate collisions.

Despite the promise of such predictive methods, the theoretical understanding of their performance and the constraints they impose remain incomplete (13). It is important to check for stability in the reinforcement learning loop, as well as to consider the overhead of collecting the observations used as inputs for the predictor network. In a high-speed context, if the velocity of a vehicle is above 70 m/s, the coherence time of the channel shrinks dramatically, and an LSTM-based predictor might supply outdated predictions by the time they are employed for contention window setting. This mismatch between prediction horizon and actual channel coherence must be quantified in design. (14)

This paper aims to address these issues and provide a framework that integrates LSTM-based temporal predictions with game-theoretic reinforcement learning for channel access. Through a series of analytic derivations, we formulate the channel access problem as a stochastic game in which each node aims to maximize its own expected throughput while preserving fairness. We then outline how deep reinforcement learning can approximate the solution of the coupled Bellman equations when nodes only partially observe the network state. Furthermore, we discuss how hidden state inference is realized within a POMDP formalism, and how collisions and channel measurements can be used to refine beliefs about the underlying state of the network. We conduct an extensive set of simulations to evaluate performance gains, highlight the operational boundaries where the predictions degrade, and point out the conditions under which learning-based MAC protocols yield minimal advantage over simpler heuristics.

The remainder of this expanded manuscript is organized as follows. We present the system model and problem formulation, clarifying the channel and mobility models used for evaluating performance (15). We then delve into the machine learning framework, including a mathematical expansion of the neural network architecture, the loss functions used for training, and the theoretical considerations regarding convergence under partial observability. Next, we describe our predictive MAC protocol design, explaining the interplay between the reinforcement learning agent and the LSTM-based predictions, with mathematical details of the contention window adaptation mechanisms. We provide results from an in-depth performance evaluation, including experiments with various node densities, speeds, and channel conditions (16). Finally, we discuss conclusions, limitations, and future research directions.

System Model and Problem Formulation

We consider a set of vehicles distributed over a coverage region that can be represented as a two-dimensional domain $\mathcal{A} \subset \mathbb{R}^2$. The time-varying number of vehicles at any slot t is denoted by $N(t)$. Each vehicle is indexed by i, for $i = 1, \ldots, N(t)$ (17). For ease of notation, we sometimes omit the dependency on t when it is clear from context. Each node can transmit over a dedicated channel of bandwidth B with transmit power P_t. The wireless channel connecting nodes i and j in slot t is modeled by a large-scale path loss component, a fading component, and shadowing. Specifically, we write

$$h_{ij}(t) = \beta_0 d_{ij}^{-\gamma} \psi_{ij}(t) \xi_{ij}(t),$$

where β_0 is a constant that encapsulates reference power at a distance of 1 meter, d_{ij} is the Euclidean distance between vehicles i and j, γ is the path loss exponent, $\psi_{ij}(t)$ represents the Nakagami-m fast fading term, and $\xi_{ij}(t)$ captures log-normal shadowing with parameters μ_{shad} and σ_{shad}. In many vehicular contexts, γ ranges from 2.0 to 4.0 depending on whether the environment is an open highway or an urban grid with multiple obstructions. (*18*)

Vehicle mobility is assumed to evolve under a modified Gauss-Markov process. Let $v_i(t)$ be the velocity of node i at slot t. The velocity is updated as (*19*)

$$v_i(t) = \alpha v_i(t-1) + (1-\alpha)\bar{v} + \sqrt{1-\alpha^2}\,\eta_i,$$

where $\bar{v}$ is the mean velocity in the region of interest, α is a memory parameter in $[0, 1]$, and η_i is sampled from a Gaussian distribution with zero mean and variance σ_v^2. Values of α close to 1 preserve the previous velocity, while smaller α leads to a quick reversion to the mean $\bar{v}$. The position of a vehicle is then updated in each slot by integrating the velocity over the slot duration. In urban settings, additional constraints or modifications can be introduced for turning at intersections or stopping at traffic lights, but these details remain conceptually consistent with the same general model.

We focus on a slotted protocol, where each timeslot is indexed by $t = 1, 2, \ldots$ (*20*). During each slot, vehicles contend for channel access. We define a general notation for the system state as

$$s(t) = \big(\mathbf{Q}(t), \mathbf{H}(t), \mathbf{V}(t)\big),$$

where $\mathbf{Q}(t)$ denotes the vector of queue lengths of all nodes at time t, $\mathbf{H}(t)$ is the matrix of channel coefficients $h_{ij}(t)$, and $\mathbf{V}(t)$ is the vector of velocities of all nodes. Other relevant components might include the unobservable interference from external sources or channel conditions from adjacent bands, but we concentrate on the key elements that define the local environment for each vehicle.

A collision occurs if more than one node transmits in the same slot within the same carrier-sense range or if the interference from simultaneous transmitters pushes the signal-to-interference-plus-noise ratio below the decoding threshold at the intended receiver (*21*). We denote the collision event for a node i by an indicator $C_i(t)$. The presence or absence of collisions is used as partial feedback in the reinforcement learning scheme.

We formulate a stochastic game $\mathcal{G} = \langle \mathcal{N}, \mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P} \rangle$, where $\mathcal{N}$ is the set of players (vehicles), $\mathcal{S}$ is the joint state space, $\mathcal{A}$ is the set of possible actions for each vehicle (such as backoff window settings or immediate transmissions), $\mathcal{R}$ is the reward function, and $\mathcal{P}$ is the state transition kernel. Each node i at state s chooses an action $a_i \in \mathcal{A}$. The collective action is denoted by $\mathbf{a} = (a_1, \ldots, a_{N(t)})$. The state evolves to a new state s' according to $\mathcal{P}(s'|s, \mathbf{a})$. Each node receives a reward $r_i(s, \mathbf{a}, s')$. The objective for node i is to find a policy π_i that maximizes the expected discounted return: (*22*)

$$\mathbb{E}\Big[\sum_{\tau=0}^{\infty} \lambda^\tau r_i(s_\tau, a_{i,\tau})\Big],$$

where $\lambda \in [0, 1)$ is the discount factor, and τ indexes future timeslots.

In many classical analyses, one seeks a Nash equilibrium in which no player unilaterally benefits from deviating from its strategy. The equilibrium can be characterized using sets of coupled Bellman equations, each describing the optimal Q-values for each node i: (*23*)

$$Q_i^{\pi_i, \pi_{-i}}(s, a_i) = \mathbb{E}\Big[r_i(s, a_i) + \lambda \sum_{s'} \mathcal{P}(s'|s, a_i, a_{-i})\, V_i^{\pi_i, \pi_{-i}}(s')\Big].$$

However, enumerating this solution space is often computationally intractable for large state or action spaces. Partial observability complicates the matter further, since each node only observes a subset of the full state. Hence, in the approach we adopt, each node runs an approximate reinforcement learning scheme with function approximation using a neural network, with the training method tailored to partial observability. (*24*)

Machine Learning Framework

The machine learning framework we consider employs neural networks to approximate the Q-function for each node while also inferring or predicting relevant components of the state for more informed decision making. We structure this approach in two interlinked parts. One part addresses real-time Q-value updates via a deep dueling Q-network (DDQN), whereas another focuses on leveraging an LSTM-based architecture to predict time-varying quantities such as contention window settings or expected network load in the upcoming slots.

Let $\mathbf{o}_t^i$ be the observation vector available to node i at time t. In the presence of partial observability, $\mathbf{o}_t^i$ may include local queue size, local collision indicators over a recent window of slots, measured SINR values for successful receptions, and any additional locally observable features (like velocity of the vehicle itself or direct messages from neighbors). The overall challenge is to use $\mathbf{o}_t^i$ to approximate the underlying global state $s(t)$. A typical approach is to incorporate a recurrent architecture so that the hidden state of the network, $\mathbf{h}_t^i$, evolves in tandem with the new observations:

$$\mathbf{h}_t^i = f_{\text{RNN}}\big(\mathbf{h}_{t-1}^i, \mathbf{o}_t^i; \theta_{\text{RNN}}\big),$$

where f_{RNN} can be an LSTM or GRU (Gated Recurrent Unit), and θ_{RNN} are trainable parameters. The hidden state

$\mathbf{h}_t^i$ then feeds into a fully connected layer or dueling head to generate Q-value estimates for the actions in $\mathcal{A}$. Concretely, if $\mathcal{A} = \{1, \ldots, A\}$, we produce

$$Q_i(\mathbf{h}_t^i, a) = V\big(\mathbf{h}_t^i; \theta_v\big) + \Big(A\big(\mathbf{h}_t^i, a; \theta_a\big) - \frac{1}{A}\sum_{a'} A\big(\mathbf{h}_t^i, a'; \theta_a\big)\Big),$$

where the trainable parameters θ_v and θ_a define the value and advantage streams, respectively (25). This decomposition often helps training stability and performance, especially in environments with large or continuous observation spaces.

In order to maintain stability when training the Q-network, we employ a target network, denoted by Q', which is periodically updated by copying the weights from the main network. This technique, introduced in seminal works on DQN, addresses the issue of the training target drifting too quickly (26). The loss function is typically expressed as:

$$\mathcal{L}(\theta) = \mathbb{E}_{(s,a,r,s')\sim\mathcal{D}}\Big[w_i\big(r + \lambda \max_{a'} Q'(s', a'; \theta') - Q(s, a; \theta)\big)^2\Big],$$

where $\mathcal{D}$ is the experience replay buffer containing past transitions, (s, a, r, s'). The factor w_i is a weight that corrects for sampling biases introduced by prioritized replay, and θ' are the parameters of the target network. When partial observability is present, we replace (s, a, r, s') with $\big(\mathbf{o}_t^i, \mathbf{h}_t^i, a_t^i, r_t^i, \mathbf{o}_{t+1}^i, \mathbf{h}_{t+1}^i\big)$ in the stored transitions, so the agent can learn the mapping from observation and hidden state to Q-values.

The second part of the framework is a predictive module that uses an LSTM-based encoder-decoder structure to forecast certain channel metrics or slot occupancy patterns (27). Let Y_t be some target variable to be predicted. This variable could represent the expected contention window size in the next timeslot, the expected collision probability, or the anticipated number of vehicles attempting to transmit in the upcoming interval. The architecture has an encoder that ingests a sequence of past observations $\{\mathbf{z}_{t-L}, \ldots, \mathbf{z}_t\}$ and produces an encoded vector $\mathbf{h}_t^{\text{enc}}$. The decoder transforms this vector into a future prediction:

$$\hat{Y}_{t+1} = \phi\big(\mathbf{h}_t^{\text{enc}}, \theta_{\text{dec}}\big),$$

where ϕ denotes the decoding function (28). Some models may add an attention mechanism that selectively focuses on relevant time steps in the past. Once the forecast $\hat{Y}_{t+1}$ is generated, it can be fed to the Q-network as part of the observation, enabling the agent to exploit knowledge of impending channel usage surges or idle periods.

Mathematically, we can formalize the LSTM update for the encoder by:

$$\begin{aligned}
\mathbf{i}_t &= \sigma\big(\mathbf{W}_i[\mathbf{z}_t \oplus \mathbf{h}_{t-1}^{\text{enc}}] + \mathbf{b}_i\big),\\
\mathbf{f}_t &= \sigma\big(\mathbf{W}_f[\mathbf{z}_t \oplus \mathbf{h}_{t-1}^{\text{enc}}] + \mathbf{b}_f\big),\\
\mathbf{o}_t &= \sigma\big(\mathbf{W}_o[\mathbf{z}_t \oplus \mathbf{h}_{t-1}^{\text{enc}}] + \mathbf{b}_o\big),\\
\tilde{\mathbf{c}}_t &= \tanh\big(\mathbf{W}_c[\mathbf{z}_t \oplus \mathbf{h}_{t-1}^{\text{enc}}] + \mathbf{b}_c\big),\\
\mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t,\\
\mathbf{h}_t^{\text{enc}} &= \mathbf{o}_t \odot \tanh(\mathbf{c}_t),
\end{aligned}$$

where σ is the sigmoid function, $\odot$ denotes element-wise multiplication, and $\oplus$ denotes concatenation of vectors (29). The decoder has a similar structure, which, starting from $\mathbf{h}_t^{\text{enc}}$, produces predicted values for the next time steps. These predictions have proven valuable in controlling the temperature parameter or bounding the range of the contention window in the MAC procedure.

Predictive MAC Protocol Design

We now describe how to integrate the two neural modules into a functioning MAC protocol that adapts in real time to the changing vehicular environment. The protocol must schedule transmissions in each slot while minimizing collisions and ensuring fairness among nodes with heterogeneous traffic demands (30). The predictive MAC protocol operates on two fundamental pillars: forecasting relevant channel characteristics for the near future, and adjusting contention or scheduling actions via reinforcement learning that includes these forecasts in the decision process.

In more standard CSMA/CA-based solutions such as IEEE 802.11p, the contention window is bounded between $CW_{\min}$ and $CW_{\max}$, and vehicles choose random backoff values within this window after each busy slot or collision. In our predictive architecture, the MAC protocol refines the selection of backoff values in real time. Denote the chosen contention window for node i at the beginning of slot t by $CW_i(t)$. A general functional form can be expressed as: (31)

$$CW_i(t) = \Big\lfloor \Omega\Big(\hat{Y}_t, Q_i(\mathbf{h}_{t-1}^i, a)\Big)\Big\rfloor,$$

where $\hat{Y}_t$ is a predicted metric (for instance, the predicted occupancy or recommended window size) from the LSTM module, and $Q_i(\mathbf{h}_{t-1}^i, a)$ is the estimated Q-value from the reinforcement learning agent's previous slot. The function Ω might be a feedforward network or a parameterized function that combines these inputs. For practical reasons, we often constrain $CW_i(t)$ to lie between $CW_{\min}$ and $CW_{\max}$, ensuring backwards compatibility and bounded delay.

A simple formulation that can be implemented in hardware-limited devices is to define:

$$CW_i(t) = \Big\lfloor CW_{\min} + (CW_{\max} - CW_{\min})\,\sigma\Big(\frac{Q_i(\mathbf{h}_{t-1}^i, a)}{T_{\text{temp}}}\Big)\Big\rfloor,$$

where σ is the sigmoid function, and T_{temp} is a temperature parameter that influences the spread of the contention window. When the Q-value is high (indicating high expected reward for transmitting immediately), the window is narrowed. Conversely, if the agent anticipates many other nodes contending or a low reward for immediate transmission, the window widens, reducing collision risk.

Furthermore, we incorporate a mechanism for intention broadcasting. Each node i periodically sends a short control packet that encodes an intention vector $\mathbf{I}_i$. This vector has length M, indicating the reservation or planned usage for the next M timeslots, with elements in $\{0, 1\}$. To construct the vector $\mathbf{I}_i$, we solve an optimization of the form:

$$\min_{\mathbf{I}_i} \left\|\mathbf{Y} - \mathbf{D}\mathbf{I}_i\right\|_2^2 + \kappa \|\mathbf{I}_i\|_1,$$

where $\mathbf{Y}$ is a target representation of the forecasted traffic demands, and $\mathbf{D}$ is an overcomplete dictionary of slot patterns. The regularization term $\|\mathbf{I}_i\|_1$ encourages sparse usage to avoid hogging the channel. This intention broadcasting approach mitigates hidden-node problems, since each node has at least partial knowledge of other nodes' intended transmissions, and can incorporate this knowledge into the Q-value estimates.

Another key design element is collision resolution. Upon detecting a collision (for instance, when an acknowledgment is not received or an explicit collision flag is broadcast), the protocol updates the reinforcement learning module with a negative reward. We define a collision penalty as $\rho < 0$. If node i experiences a collision in slot t, then $r_i(t) = \rho$ (32). If node i transmits successfully, it can receive a positive reward proportional to the throughput or the number of bits successfully delivered. If node i refrains from transmitting (idle action) and does not experience a collision, the reward might be a small penalty or zero. The exact structure of the reward function significantly affects system behavior; for example, weighting collisions heavily fosters more conservative backoff expansions, while emphasizing successful transmissions encourages the agent to transmit more aggressively. (33)

In partial observability, each node must also estimate how many of its neighbors are currently in backoff and how many are waiting for channel access. The observations of collisions and idle slots, combined with local channel measurements such as the measured channel power during idle slots, inform a Bayesian update of each node's belief state. A simplified approach might rely purely on recurrent neural networks to maintain a hidden state that effectively represents these beliefs (34). Specifically, each node's hidden state evolves based on the immediate feedback in slot t:

$$\mathbf{h}_t^i = R_\theta\big(\mathbf{h}_{t-1}^i, C_i(t), I_i(t), \ldots\big),$$

where $I_i(t)$ is an indicator for successful idle detection, and R_θ is a learned recurrent function. This hidden state, once updated, is used to calculate Q-values for the next slot. In principle, more sophisticated factoring of the state space can be considered, but in practice, a well-structured recurrent network or attention mechanism can suffice to capture the relevant features.

Performance Evaluation

We now present a comprehensive performance evaluation that assesses how the proposed predictive MAC protocol performs under various conditions (35). In the interest of capturing realistic vehicular environments, we conduct simulations using the OMNeT++ simulator, integrated with the Veins framework, which can model both the wireless channel and vehicular mobility. Our experiments span multiple scenarios, including a 500-meter highway segment and a Manhattan-style urban grid. We vary node density from as low as 20 vehicles up to 200 vehicles to test scalability, and we examine different mobility speeds ranging from 30 km/h to 200 km/h. (36)

In each simulation run, we measure several key metrics. One primary metric is access delay, defined as the time from when a packet arrives at the MAC queue until it is either successfully transmitted or dropped. Another important metric is packet delivery ratio (PDR), which is the fraction of transmitted data packets that are successfully received by their intended recipients (37). Collision rate is monitored to measure how well the reinforcement learning scheme balances aggressiveness and caution in contending for the channel. We also examine fairness metrics, often using a Jain's fairness index, to evaluate how equitably resources are shared among vehicles with different traffic demands.

Simulation results highlight that under moderate mobility speeds (up to about 70 m/s), the proposed protocol achieves substantially lower delay compared to 802.11p (38). Specifically, in a highway scenario with 100 vehicles, the median access delay of 802.11p was approximately 4.7 ms, while the predictive MAC approach reduced this to around 3.1 ms. A deeper inspection reveals that our method is able to predict bursts of transmissions triggered by multiple vehicles entering the same region, thus expanding the contention window proactively. Conversely, during periods of idle or low contention, the protocol narrows the window, taking advantage of idle slots to transmit quickly. These features combine to offer more efficient channel utilization. (39)

However, the advantages of the learning-based protocol erode at very high mobility, such as speeds exceeding 140 km/h. In those cases, our LSTM module produces predictions that are quickly invalidated by the rapid changes in vehicle positions and channel conditions. A typical manifestation is a mismatch between the predicted collision probability and the actual collision probability, causing suboptimal backoff decisions (40). This effect can cause the delay to rise by around 72 percent compared to the performance at moderate speeds. In some scenarios, the performance under these extreme speeds approaches that of simpler protocols, which indicates that the cost of performing predictions may not be justified if the environment changes too quickly.

Another interesting aspect is the effect of non-line-of-sight (NLOS) conditions in tunnels or heavily shadowed city streets (*41*). Our approach relies substantially on collision measurements and partial CSI estimates in consecutive slots. When signals are heavily attenuated or the environment is extremely dynamic, the predicted backlog or collision probability might become inaccurate. In tunnel scenarios with deep fading, the measured PDR for our protocol dropped to about 68 percent, which is similar to that of legacy systems. This suggests that advanced solutions, such as cooperative relaying or dedicated roadside units that relay real-time channel occupancy data, may be necessary to sustain predictive performance in extreme NLOS conditions.

Regarding the training process, we measured the convergence of the reinforcement learning agent in terms of the average per-slot reward, as well as the fraction of time the system remains in a high-collision state. In typical runs with 100 vehicles, the agent starts from a random policy and experiences a high collision rate initially. As the number of training episodes increases, the collision rate steadily declines until about 3,500 episodes, by which time the Q-values have stabilized sufficiently to yield a near-optimal policy (*42*). Additional training after about 6,000 episodes yields diminishing returns, with the QoS violation probability plateauing around 0.12. This relatively slow training speed is largely due to partial observability, where each node must rely on sparse collision feedback and local SINR measurements. Techniques like prioritized replay and the dueling architecture do help accelerate convergence. (*43*)

From an energy consumption standpoint, we compared the proposed scheme against a TDMA-based approach that forces all nodes to remain synchronized in a fixed schedule. Although TDMA can guarantee collision-free operation, it often forces nodes to transition from sleep to awake states more frequently to maintain timing synchronization. Our approach, by learning to concentrate transmissions when beneficial and allow longer idle periods otherwise, achieves about a 23 percent reduction in the total number of radio wake-ups. This translates to a significant saving in energy usage, which is a non-trivial advantage for battery-electric or hybrid vehicles that must conserve power for propulsion, sensing, and onboard computing.

Lastly, an important question concerns the computational overhead of running the neural networks on embedded hardware in vehicles. In our experiments, the neural network inference and LSTM predictions were assumed to be computed on a standard CPU with sufficient floating-point capability. Nonetheless, in practice, specialized accelerators or GPU-based solutions might be necessary to meet strict sub-millisecond deadlines. The overhead from each inference must be carefully evaluated to ensure that the latency introduced by the neural modules does not undermine the benefit of adaptation. Preliminary hardware-in-the-loop tests indicate that optimized implementations of LSTMs can run within a fraction of a millisecond on modern embedded accelerators, but further research is needed to confirm these results under cost and power constraints typical of automotive environments.

Conclusion

This work demonstrates that an integrated machine learning MAC protocol can deliver substantial improvements in delay, throughput, and collision reduction for vehicular networks operating under moderate speeds and typical traffic conditions (*44*). By modeling the channel access process as a partially observable Markov decision process and pairing a dueling Q-network with an LSTM predictor, it becomes possible to exploit spatial-temporal correlations in a constantly shifting vehicular scenario. Our results indicate a notable improvement in performance compared to baseline IEEE 802.11p, particularly in environments where bursty and correlated transmissions are significant.

Despite these promising outcomes, there are limitations that must be acknowledged. When mobility is extremely high, the prediction horizon of the LSTM is often reduced to a point where its benefits vanish. The difficulties inherent in training under partial observability, combined with the short coherence time of the channel, mean that the agent's learned policies can degrade sharply outside of their training conditions. Additionally, in heavily shadowed or NLOS scenarios, local channel measurements provide incomplete information about the environment, constraining the learning-based protocol to operate at a level similar to that of legacy techniques. We also note that real hardware implementations require specialized accelerators to meet sub-millisecond decisions. (*45*)

Future research directions could explore federated learning or distributed training, whereby groups of vehicles share neural network updates to accelerate convergence and reduce the training overhead at each node. Another avenue of improvement is in hybrid solutions that combine model-based predictions for large-scale mobility patterns with data-driven neural networks for finer-scale channel fluctuations. Such hybrid solutions might generalize better to unforeseen environments, as they preserve some degree of explicit modeling for macroscopic processes, while still leveraging deep learning for local channel adaptation (*46*). Integrating side information from high-definition maps or infrastructure-based sensors could further refine predictions of traffic patterns and channel conditions, leading to a more robust protocol.

The proposed predictive MAC architecture underscores the potential of combining machine learning with advanced channel modeling to tackle the unique challenges of next-generation vehicular networking. By carefully balancing real-time predictions with reinforcement learning policies, vehicles can achieve proactive, data-driven channel access decisions that maximize throughput and fairness while respecting latency constraints. With ongoing innovations in both the theoretical and practical domains, there is a clear trajectory toward more adaptive, intelligent MAC protocols that can handle the evolving demands of connected and autonomous vehicles. (47–52)

References

1. Gaines, B. R. The learning of perceptual-motor skills by men and machines and its relationship to training. *Instructional Science*, Vol. 1, No. 3, 1972, pp. 263–312. doi:10.1007/bf00055704.
2. Schirmer, C. M. Inferior Vena Cava Distensibility in Patients with SAH, New Technology and Numbers, Better Care? *Neurocritical care*, Vol. 13, No. 1, 2010, pp. 1–2. doi:10.1007/s12028-010-9426-2;10.1007/s12028-010-9378-6.
3. Pequegnat, W., B. R. S. Rosser, A. M. Bowen, S. Bull, R. J. DiClemente, W. O. Bockting, J. Elford, M. Fishbein, L. J. Gurak, K. J. Horvath, J. A. Konstan, S. M. Noar, M. W. Ross, L. Sherr, D. Spiegel, and R. S. Zimmerman. Conducting Internet-based HIV/STD prevention survey research: considerations in design and evaluation. *AIDS and behavior*, Vol. 11, No. 4, 2006, pp. 505–521. doi:10.1007/s10461-006-9172-9.
4. Ahmed, M. M., A. Ganguly, A. Vashist, and S. M. P. Dinakarrao. AWARe-Wi: A jamming-aware reconfigurable wireless interconnection using adversarial learning for multichip systems. *Sustainable Computing: Informatics and Systems*, Vol. 29, 2021, pp. 100470–. doi:10.1016/j.suscom.2020.100470.
5. null Sam, C. Pvs, Thomson, null Gopika, H. Smith, J. Chan, Kumar, Field, Salmasi, null Yousuf, S. Pirola, null Sasidharan, null Sumesh, S. M. Fisichella, null Jarral, null O'regan, null Moore, X. Xu, Y. Xiao, null Athanasiou, S. Alwis, D. Mozalbat, Cyclewala, null Athansiou, M. Yousuf, Nienabar, U. Simoniuk, M. Ntouskou, M. S. Shaw, Richards, Muneer, Kuduvalli, Field, null Theologu, Y. M. Oo, M. Hossack, Fisher, F. Torella, M. Field, Madine, and null Akhtar. SCTS Annual Meeting 2022 Abstracts. *Journal of cardiothoracic surgery*, Vol. 18, No. Suppl 1, 2023, pp. 198–. doi:10.1186/s13019-023-02292-4.
6. Iskandarani, M. Z. Vehicular Connectivity Analysis Using Enhanced Quality Slotted ALOHA (EQS-ALOHA). In *Science and Information Conference*. Springer, 2024, pp. 484–509.
7. Petkov, V. and K. Obraczka. CCNC - The case for using traffic forecasting in schedule-based channel access. In *2011 IEEE Consumer Communications and Networking Conference (CCNC)*. IEEE, 2011, pp. 208–212. doi:10.1109/ccnc.2011.5766456.
8. Dhakal, S., Q. Chen, D. Qu, D. Carillo, Q. Yang, and S. Fu. Sniffer faster r-cnn: A joint camera-lidar object detection framework with proposal refinement. In *2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST)*. IEEE, 2023, pp. 1–10.
9. Land, S. M. and M. J. Hannafin. Patterns of understanding with open-ended learning environments: A qualitative study. *Educational Technology Research and Development*, Vol. 45, No. 2, 1997, pp. 47–73. doi:10.1007/bf02299524.
10. Enot, D., W. Lin, M. Beckmann, D. A. Parker, D. P. Overy, and J. Draper. Preprocessing, classification modeling and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data. *Nature protocols*, Vol. 3, No. 3, 2008, pp. 446–470. doi:10.1038/nprot.2007.511.
11. Kandasamy, K., S. Srinivas, K. Achuthan, and V. P. Rangan. IoT cyber risk: a holistic analysis of cyber risk assessment frameworks, risk vectors, and risk ranking process. *EURASIP Journal on Information Security*, Vol. 2020, No. 1, 2020, pp. 1–18. doi:10.1186/s13635-020-00111-0.
12. MacCallum, M. Computer algebra in gravity research. *Living reviews in relativity*, Vol. 21, No. 1, 2018, pp. 6–6. doi:10.1007/s41114-018-0015-6.
13. Bontozoglou, A., K. Yang, and K. Guild. A midterm DBA algorithm for quality of service on aggregation layer EPON networks. *Photonic Network Communications*, Vol. 25, No. 2, 2013, pp. 120–134. doi:10.1007/s11107-013-0396-0.
14. Wandell, R., M. S. Hossain, and I. Hussain. A cost-effective Wi-Fi-based indoor positioning system for mobile phones. *Wireless Networks*, Vol. 29, No. 6, 2023, pp. 2845–2862. doi:10.1007/s11276-023-03362-0.
15. Wadsworth, J. Forging the Solution to the Energy Challenge: The Role of Materials Science and Materials Scientists. *Metallurgical and Materials Transactions B*, Vol. 41, No. 2, 2010, pp. 259–274. doi:10.1007/s11663-010-9349-5.
16. Heuston, E. F., C. A. Keller, J. Lichtenberg, B. Giardine, S. M. Anderson, R. C. Hardison, and D. M. Bodine. Establishment of regulatory elements during erythro-megakaryopoiesis identifies hematopoietic lineage-commitment points. *Epigenetics & chromatin*, Vol. 11, No. 1, 2018, pp. 22–22. doi:10.1186/s13072-018-0195-z.
17. Amalvy, A., M. Janickyj, S. Mannion, P. MacCarron, and V. Labatut. Interconnected Kingdoms: comparing 'A Song of Ice and Fire' adaptations across media using complex networks. *Social Network Analysis and Mining*, Vol. 14, No. 1. doi:10.1007/s13278-024-01365-z.
18. Bell, E. Y., K. Ogata, M. T. Lucas, W. B. Lees, L. M. Jackson, J. M. Mitchem, B. J. Little, C. F. Amer, C. Weber, D. L. Hardesty, J. Wanser, L. Wurst, J. H. Jameson, J. Wanser, M. A. Pfeiffer, W. B. Barr, C. T. Espenshade, L. E. Babits, D. C. Linck, and D. W. Babson. Reviews. *Historical Archaeology*, Vol. 29, No. 4, 1995, pp. 84–114. doi:10.1007/bf03374218.
19. Volodko, A., E. Cockbain, and B. Kleinberg. "Spotting the signs" of trafficking recruitment online: exploring the characteristics of advertisements targeted at migrant job-seekers. *Trends in Organized Crime*, Vol. 23, No. 1, 2019, pp. 7–35. doi:10.1007/s12117-019-09376-5.

20. Sun, J. and L. Tesfatsion. Dynamic Testing of Wholesale Power Market Designs: An Open-Source Agent-Based Framework. *Computational Economics*, Vol. 30, No. 3, 2007, pp. 291–327. doi:10.1007/s10614-007-9095-1.
21. Zarrin, J., H. W. Phang, L. Saheer, and B. Zarrin. Blockchain for decentralization of internet: prospects, trends, and challenges. *Cluster computing*, Vol. 24, No. 4, 2021, pp. 1–26. doi: 10.1007/s10586-021-03301-8.
22. Lavarch, M., M. Sharpe, S. Webb, W. S. Young, G. O'Keefe, S. Gong, H. Tochon-Danguy, D. Gunawardana, M. Wada, K. Rykers, A. M. Scott, O. Acosta, R. Li, H. Frimmel, M. Caon, D. L. Collins, I. G. McLean, and S. Ourselin. EPMS'06 conference abstracts. *Australasian Physics & Engineering Sciences in Medicine*, Vol. 29, No. 4, 2006, pp. 333–414. doi: 10.1007/bf03178400.
23. Al-Tous, H., T. Ponnada, C. Studer, and O. Tirkkonen. Multipoint channel charting-based radio resource management for V2V communications. *EURASIP Journal on Wireless Communications and Networking*, Vol. 2020, No. 1, 2020, pp. 1–20. doi:10.1186/s13638-020-01723-3.
24. Szczepura, A., P. Saravanan, D. Baines, S. Petrou, R. Crossman, A. Hardy, N. T. Raymond, A. Wilson, J. P. O'Hare, and G. Iccds. Cost-effectiveness of intermediate care clinics for diabetes (ICCD): findings of a UK cluster randomised trial. *Diabetologia*, Vol. 55, No. S1, 2012, pp. 207–207. doi:10.1007/s00125-012-2688-9.
25. Lee, S., K. Grover, and A. Lim. Enabling actionable analytics for mobile devices: performance issues of distributed analytics on Hadoop mobile clusters. *Journal of Cloud Computing: Advances, Systems and Applications*, Vol. 2, No. 1, 2013, pp. 15–. doi:10.1186/2192-113x-2-15.
26. Hwang, S. O., S. Auephanwiriyakul, M. U. Akram, B.-M. Goi, and C. S. Chan. Special issue on "Green and Human Information Technology 2019". *Neural Computing and Applications*, Vol. 32, No. 17, 2020, pp. 13077–13078. doi: 10.1007/s00521-020-04949-w.
27. Huntley, J. Teaching Milton by Computer. *Journal of Computing in Higher Education*, Vol. 3, No. 1, 1991, pp. 62–84. doi:10.1007/bf02942598.
28. Cheng, Z., D. Devarajan, and R. J. Radke. Determining vision graphs for distributed camera networks using feature digests. *EURASIP Journal on Advances in Signal Processing*, Vol. 2007, No. 1, 2006, pp. 220–220. doi:10.1155/2007/57034.
29. Yao, K., Y. Xu, H. Li, J. Chen, S. Zhang, and X. Yu. Leveraging partially overlapping channels for intra- and inter-coalition communication in cooperative UAV swarms. *Science China Information Sciences*, Vol. 64, No. 4, 2021, pp. 1–16. doi: 10.1007/s11432-020-3012-3.
30. Abdel-Fattah, M., A. Mostafa, D. Young, and I. Ramsay. Impact of transobturator tension free vaginal tapes on quality of life and sexual function in women with mixed urinary incontinence. *International urogynecology journal*, Vol. 21, No. S1, 2010, pp. S1–428. doi:10.1007/s00192-010-1192-3.
31. Huang, H., Y. Zhang, V. Gallegos, N. Sorrelle, M. Zaid, J. E. Toombs, W. Du, S. M. Wright, M. M. Hagopian, Z. Wang, A. N. Hosein, A. A. Sathe, C. Xing, E. J. Koay, K. Driscoll, and R. A. Brekken. Targeting TGFR2-mutant tumors exposes vulnerabilities to stromal TGF blockade in pancreatic cancer. *EMBO molecular medicine*, Vol. 11, No. 11, 2019, pp. e10515–. doi:10.15252/emmm.201910515.
32. Mitchell, A., T. J. Welsh, M. C. Watson, J. Snowball, and A. McGrogan. P-509 Oral anticoagulant prescribing in people aged ¿= 75 years with atrial fibrillation in the UK pre- and post-introduction of direct oral anticoagulants: a study protocol. *European Geriatric Medicine*, Vol. 10, No. S1, 2019, pp. 1–325. doi:10.1007/s41999-019-00221-0.
33. Windschitl, M. Instructional Animations: The In-House Production of Biology Software. *Journal of Computing in Higher Education*, Vol. 7, No. 2, 1996, pp. 78–94. doi:10.1007/bf02948595.
34. Rosen, E. F. and L. C. Petty. Selecting an appropriate multimedia authoring language. *Behavior Research Methods, Instruments, & Computers*, Vol. 28, No. 2, 1996, pp. 286–290. doi:10.3758/bf03204783.
35. Qadir, J. Artificial Intelligence Based Cognitive Routing for Cognitive Radio Networks. *Artificial Intelligence Review*, Vol. 45, No. 1, 2015, pp. 25–96. doi:10.1007/s10462-015-9438-6.
36. Ogaji, J. Sustainable Agriculture in the UK. *Environment, Development and Sustainability*, Vol. 7, No. 2, 2005, pp. 253–270. doi:10.1007/s10668-005-7315-1.
37. Cajander, S., B. S. Jaliff, A. Somell, A. Bäckman, H. Alpkvist, V. Özenci, J. Wernerman, and K. Strålin. HLA-DRA and CD74 on intensive care unit admission related to outcome in sepsis. *Critical care (London, England)*, Vol. 22, No. Suppl 1, 2018, pp. 183–184. doi:10.1186/s13054-018-1973-5.
38. Koller, J. R., D. A. Jacobs, D. P. Ferris, and C. D. Remy. Learning to walk with an adaptive gain proportional myoelectric controller for a robotic ankle exoskeleton. *Journal of neuroengineering and rehabilitation*, Vol. 12, No. 1, 2015, pp. 97–97. doi:10.1186/s12984-015-0086-5.
39. Chin, T., K. Xiong, and E. Blasch. CRAMStrack: Enhanced Nonlinear RSSI Tracking by Using Circular Multi-Sectors. *Journal of Signal Processing Systems*, Vol. 93, No. 1, 2020, pp. 79–97. doi:10.1007/s11265-020-01516-3.
40. Sherman, K. A., D. Harcourt, L.-K. Shaw, C. J. Winch, P. Brown, L. D. Cameron, T. Lam, and J. Boyages. The Breconda trial: facilitating surgical decision making of women with breast cancer. *International journal of behavioral medicine*, Vol. 21, No. S1, 2014, pp. 0465–154. doi:10.1007/s12529-014-9418-2.
41. Ozik, J., N. Collier, J. M. Wozniak, C. M. Macal, C. Cockrell, S. H. Friedman, A. Ghaffarizadeh, R. Heiland, G. An, and P. Macklin. High-throughput cancer hypothesis testing with an integrated PhysiCell-EMEWS workflow. *BMC bioinformatics*, Vol. 19, No. 18, 2018, pp. 81–97. doi:10.1186/s12859-018-2510-x.

42. Kyle, K., S. A. Crossley, and C. M. Berger. The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0. *Behavior research methods*, Vol. 50, No. 3, 2017, pp. 1030–1046. doi:10.3758/s13428-017-0924-4.
43. Koo, S. G. M., C. Rosenberg, H.-H. Chan, and Y. C. Lee. Location discovery in enterprise-based wireless networks: Case studies and applications. *Annales Des Télécommunications*, Vol. 58, No. 3, 2003, pp. 531–552. doi:10.1007/bf03001027.
44. Turner, D., B. Gray, S. D. Luzio, S. C. Bain, G. Dunseath, M. Campbell, D. J. West, and R. M. Bracken. Resistance exercise induces increases in circulating interleukin-6 in type 1 diabetes individuals. *Diabetologia*, Vol. 56, No. 1, 2013, pp. 1–566. doi:10.1007/s00125-013-3012-z.
45. Keller, F., S. Gunasekharan, N. Mayo, and M. Corley. Timing accuracy of Web experiments: A case study using the WebExp software package. *Behavior research methods*, Vol. 41, No. 1, 2009, pp. 1–12. doi:10.3758/brm.41.1.12.
46. Helbing, D. and S. Balietti. From social data mining to forecasting socio-economic crises. *The European physical journal. Special topics*, Vol. 195, No. 1, 2011, pp. 3–68. doi: 10.1140/epjst/e2011-01401-8.
47. Juarez, E. F., R. Lau, S. H. Friedman, A. Ghaffarizadeh, E. A. Jonckheere, D. B. Agus, S. M. Mumenthaler, and P. Macklin. Quantifying differences in cell line population dynamics using CellPD. *BMC systems biology*, Vol. 10, No. 1, 2016, pp. 92–92. doi:10.1186/s12918-016-0337-5.
48. Dhakal, S., D. Qu, D. Carrillo, Q. Yang, and S. Fu. Oasd: An open approach to self-driving vehicle. In *2021 Fourth International Conference on Connected and Autonomous Driving (MetroCAD)*. IEEE, 2021, pp. 54–61.
49. Belsley, D. A. Mathematica as an Environment for Doing Economics andEconometrics. *Computational Economics*, Vol. 14, No. 1, 1999, pp. 69–87. doi:10.1023/a:1008699629392.
50. Alhomayani, F. and M. H. Mahoor. OutFin, a multi-device and multi-modal dataset for outdoor localization based on the fingerprinting approach. *Scientific data*, Vol. 8, No. 1, 2021, pp. 1–14. doi:10.1038/s41597-021-00832-y.
51. Yang, Y., M. Fu, M.-D. Li, K. Zhang, B. Zhang, S. Wang, Y. Liu, W. Ni, Q. Ong, J. Mi, and X. Yang. O-GlcNAc transferase inhibits visceral fat lipolysis and promotes diet-induced obesity. *Nature communications*, Vol. 11, No. 1, 2020, pp. 181–181. doi:10.1038/s41467-019-13914-8.
52. Gryde, S. K. The feasibility of “programed” television instruction. *AV communication review*, Vol. 14, No. 1, 1966, pp. 71–89. doi:10.1007/bf02768509.