RESEARCH PAPER

MEASURING MOBILITY AND LINK REALIABILITY USING CSS AND MACHINE LEARNING METHODS IN CR-VANET UTILIZING GRADIENT BOOSTING MACHINE (GBM)

Gaurav Gupta[1], Dr. Amit Dixit[2], Dr. Nishant Saxena[3]

[1]Ph.D Scholar (CSE), Quantum University, Roorkee. Email: gaurav4584ster@gmail.com | Contact No: 8126834983

[2]Professor ECE, Quantum University, Roorkee. Email: amitdixit.ece@quantumeducation.in | Contact No: 8650505814

[3]Professor ECE, Tula's Institute, Dehradun, Uttarakhand. Email: nishant_ei@rediffmail.com | Contact No: 9411945419

ABSTRACT

Cooperative Spectrum Sensing, Cognitive Radio, VANET, Gradient Boosting Machine, Link Reliability, Machine Learning. The current study provides a Gradient Boosting Machine-based model for carrying out estimations of reliability and mobility that utilizes Cooperative Spectrum Sensing and machine learning to boost the flexibility and robustness of vehicular communication. In order to estimate reliability while considering various vehicular and spectrum conditions, this model utilizes a wide variety of characteristics, including vehicle speed, Received Signal Strength, Signal-to-Noise Ratio, and local sensing decisions. The GBM algorithm, which is a combination of several weak classifiers, also provides better prediction accuracy compared to traditional classifiers like Decision Tree (DT) and Random Forest (RF) in terms of nonlinear relationships between the characteristics of mobility and communication. Simulation results using the SUMO and NS-3 scenarios have demonstrated the effectiveness of the proposed GBM model with a higher accuracy of 94.6%, a lower RMSE of 0.042, and a higher ROC-AUC of 0.96. Throughput, link reliability, and Quality of Service (QoS) may all be improved by combining CSS with ML-based prediction to enable more intelligent spectrum decision-making. In order to support autonomous driving, safety-critical communication, and real-time spectrum awareness, our study creates a scalable basis for 6G-enabled intelligent automotive systems. Future research will concentrate on cross-layer optimization, federated learning integration, and real-world validation to enable adaptive and energy-efficient CR-VANET communication.

Keywords: Gradient Boosting Machine, Cognitive Radio, VANET, Cooperative Spectrum Sensing, Link Reliability, Machine Learning, Mobility Prediction.

How to cite this article: Gupta G, Dixit A, Saxena N. Measuring Mobility and Link Reliability Using CSS and Machine Learning Methods in CR-VANET Utilizing Gradient Boosting Machine (GBM). Int J Drug Deliv Technol. 2026;16(56s): 1304-1315. DOI: 10.25258/ijddt.16.56s.145

Source of support: Nil.

Conflict of interest: None.

SECTION I

INTRODUCTION

Reliable, high-capacity vehicular communication is in high demand due to the exponential expansion of connected and autonomous cars. Cellular networks and traditional dedicated short-range communication (DSRC) are unable to effectively manage dynamic, location-dependent spectrum scarcity. A potential paradigm that allows secondary users (SUs) to opportunistically access licensed bands without interfering with main users (PUs) is called cognitive radio (CR) [1],[2]. Vehicles may dynamically take advantage of spectrum possibilities by integrating CR with Vehicular Ad-hoc Networks (VANETs) to create Cognitive Radio VANETs (CR-VANETs). This greatly improves bandwidth usage and communication dependability in intelligent transportation systems.[3],[4].

Spectrum sensing, which identifies channel availability by detecting PU activity, is a basic procedure in CR-VANETs. However, fading, shadowing, and noise uncertainty are problems with single-node sensing that are made worse by ambient heterogeneity and vehicle motion. Cooperative Spectrum Sensing (CSS) combines sensing information from several nodes to detect PUs more precisely in order to remedy this.[5],[6]. However, CSS is faced with challenges such as latency in reports, instability of control connections, and spatial irregularities in fast-moving vehicles [7][8].

In addition, the high vehicle mobility reduces the reliability of the link due to the dynamic topology of the network and the fast signal fluctuations, which decrease the lifespan of the connections as well as the decisions made by the sensor. The signal to noise ratio changes as well as the disconnections reduce the sensor transmission rates. Reliable mobility and link estimation techniques are therefore essential for proactive spectrum access and consistent communication in CR-VANETs [9], [10].

The application of machine learning (ML) to link-quality prediction and spectrum sensing has expanded significantly. Data-driven modelling is used by ML to improve decision accuracy in the face of uncertainty [11], [12].Nonlinear connections between channel characteristics and PU activity have been extracted using deep learning architectures like CNNs and LSTMs [13].

Lightweight alternatives that can effectively infer at the network edge are motivated by the fact that these models are computationally costly and inappropriate for low-latency automotive devices [14].

Gradient Boosting Machines (GBMs), such as XGBoost, LightGBM, and CatBoost, provide a convincing trade-off between accuracy, interpretability, and computational cost among ensemble machine learning techniques [15]. Tabular, heterogeneous information including received signal strength, SNR, mobility indicators (speed, direction change, distance), and packet-loss statistics are well handled by GBMs [16], [17]. GBM models have proven to be highly predictive in traffic prediction, signal-quality assessment, and routing optimization in wireless and vehicular systems [18], [19].

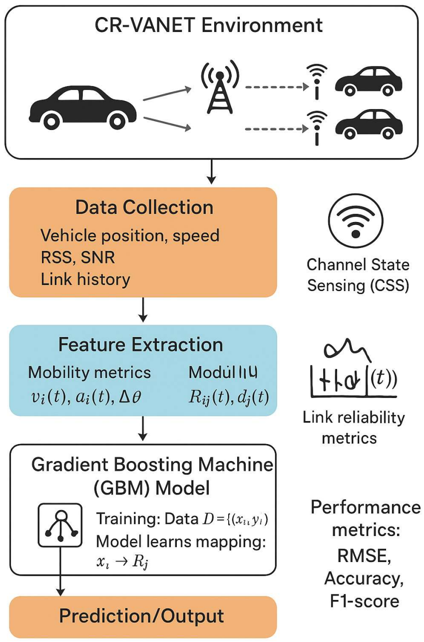


Fig1. Process for measuring mobility and link reliability using CSS and machine learning methods in CR-VANET utilizing Graidiant Boosting Machine(GBM)

There are still few integrated systems that combine CSS outputs with mobility-aware GBM-based link reliability prediction for CR-VANETs, despite advances in CR and ML. The majority of previous studies either (i) use traditional statistical models that exclude mobility characteristics or (ii) depend on deep networks that are expensive and hard to train [20]. In order to close this gap, this research suggests a mobility-aware GBM framework that predicts both short-term channel availability and link dependability by combining cooperative sensor outputs with vehicular dynamics. In highly mobile CR-VANET contexts, the suggested approach facilitates proactive spectrum access, lessens interference to PUs, and enhances link reliability and routing efficiency.

Problem Statement

The major issue this project seeks to address is:
How can cooperative spectrum sensing and machine learning techniques be employed to estimate the reliability of a vehicular communication link, as defined by the probability of continuing connectivity, under a CR-VANET environment, with consideration of the effects of high mobility, channel conditions, and spectrum availability?

Key challenges include:
High mobility and topology changes: Vehicles tend to have high mobility rates, which often leads to disconnections or unpredictable connection intervals.
Spectrum dynamics and uncertainty: In the CR environment, channel availability is highly dynamic and unpredictable, with the results of cooperative sensing possibly being unclear or delayed.
Complex relationships between the mobility features (speed, relative distance), neighbour density, channel occupancy, as well as link quality metrics such as SNR and RSSI, exhibit non-linear relationships.
Real-time requirements: For the purposes of routing or link selection decisions in CR-VANETs, the accuracy, efficiency, and generality of the link reliability estimation should be high.

Research Contributions:

For the purpose of estimating the mobility and reliability of the links in CR-VANETs, the current article uses a

Gradient Boosting Machine method, which incorporates mobility/link quality feature data and CSS data.
Among the main contributions of the current article, the following may be listed:

Integration of Multi-Source Features:
We gather data from various sources of vehicle mobility (speed, distance, neighbor density), channel/link quality (SNR, RSSI), as well as cooperative spectrum sensing (channel occupancy, detection/false alarm probability). This provides a complete feature vector as input to the reliability prediction model**.**

Machine-learning based reliability estimator:
We have chosen the GBM approach as it offers non-linear modeling, noise tolerance, as well as feature importance analysis. We compare it with basic approaches based on individual Decision Trees as estimators.

Performance Evaluation & Comparison:
We simulate the CR-VANET environment, evaluate the superiority of the GBM model over basic approaches in terms of metrics (accuracy, F1-score, ROC/AUC), as well as feature importance analysis to evaluate the parameters contributing maximally to the reliability of the links.

Implications of the model in routing/link selection:
The model offers a probabilistic reliability score of the links, allowing better route/link selection in CR-VANETs by addressing the dynamic interaction between spectrum resources & vehicle mobility.

Paper Organization:
The rest of the paper is organized as follows: In Section II, the literature on CR-VANETs, prediction of mobility, cooperative spectrum sensing, and link reliability estimation using ML techniques is extensively discussed. In Section III, the problem formulation, feature set, and system model have been discussed. In Section IV, the training process and proposed GBM-based estimation technique have been discussed. In Section V, the model comparison, results, and simulation have been depicted. In Section VI, the limitations of the paper have been discussed, along with the future work. Finally, the investigation is completed in Section VII.

Novelty of the Proposed Framework

Though there have been some works on the applications of Gradient Boosting Machine (GBM) for networking and wireless communication, the proposed work is unique and has an integrated framework especially designed for Cognitive Radio Vehicular Ad-hoc Networks (CR-VANETs). The novelty of the proposed work is the integration of cooperative spectrum sensing, vehicular mobility, and wireless communication quality into a single machine learning model. The innovative features of the proposed work are as follows:

A GBM-based wireless communication link reliability prediction model for CR-VANET environments, especially for the integration of mobility.
Integration of mobility and communication quality along with cooperative spectrum sensing.
A single machine learning model for proactive routing and spectrum access decisions, especially for predicting wireless communication link reliability.
Hybrid SUMO and NS3 simulation environment for validating the proposed model, especially for simulating realistic vehicular communication environments.

The proposed model for spectrum usage and wireless communication link reliability prediction, especially for the upcoming intelligent transportation systems, is based on a more reliable and intelligent decision-making framework.

SECTION II
LITERATURE REVIEW

Cognitive Radio in VANETs (CR-VANETs)
In past research on CR, it was demonstrated how CR is useful for vehicle networks, but it also identified some challenges associated with CR, such as the need for cooperative techniques with low latency, high mobility, and fast channel reconfiguration. As challenges, these works also address spectrum management, security, and deployment constraints.[21]. Recent comprehensive evaluations of cognitive radio and its connection to vehicular networks emphasize the rising significance of AI/ML strategies to control nonstationary spectrum situations in CR-VANETs**.** [22].

Key takeaways:

Although CR offers spectrum agility, it necessitates precise spectrum detection and coordination between mobile nodes [21].
In order to integrate CR into VANETs, mobility-aware spectrum access techniques, PU protection, and cooperative decision fusion must be addressed [23].

Cooperative Spectrum Sensing (CSS)

A key component of CR systems is cooperative spectrum sensing (CSS), which is especially crucial in automotive situations when individual sensors have extremely fluctuating SNRs and sporadic observations. While ML and deep-learning techniques have lately gained popularity to enhance detection under fading, noise uncertainty, and non-Gaussian interference, traditional CSS fusion rules (OR/AND/Voting, likelihood ratio tests) are still helpful. In mobile and multiband environments, multifeatured and learning-based fusion techniques (CNNs, attention networks, and GMMs) have demonstrated enhanced resilience [24].

Representative recent works:

Neural network fusion and multifeatured for CSS that enhances detection under various signal kinds [25].

To manage heterogeneous PU signals and mobility, deep and attention-based CSS is used for multiband vehicular sensing [26].

Obstacles mentioned in the literature:

CSS reporting overhead and latency in fast-moving vehicles [27].
Resilience to hostile actors and tainted reports (trust and security in CSS) [28].

Predicting Mobility and Estimating Link Lifetime in VANETs

Because mobility dynamics have a significant impact on route stability and connection dependability, predicting vehicle mobility and link lifespan is a prominent area of study interest. Techniques include statistical and machine learning models (SVM, Random Forests, SVR, XGBoost, neural nets) and physics/kinematic models (relative speed and geometry). In order to prevent leakage, recent empirical research supports time-aware validation and data-driven link-lifetime prediction [29].

Key points:

For link lifespan and mobility prediction in diverse traffic circumstances, machine learning models (such as XGBoost and GBM variants) frequently perform better than linear/statistical baselines [29].

Relative distance, speed, and direction-based Link Lifetime (ELD) models offer comprehensible baselines and perform well in straightforward situations [30].

Machine-Learning for Link Quality / Reliability Estimation:

There has been a surge in machine learning techniques for link-quality estimation (LQE) in wireless networks. Empirical research and surveys demonstrate that machine learning models are more effective than threshold-based techniques in capturing the nonlinear correlations between packet delivery/link reliability and SNR/RSSI/LQI/time-series characteristics. Because class imbalance and decision costs vary by application, community best practices include cautious feature engineering (temporal windows, moving averages), cross-scenario validation, and providing numerous metrics (ROC-AUC, PR curves) [31].

Typical ML techniques used for LQE in wireless and automotive settings:

Random forests and decision trees provide as quick, understandable baselines [32].
High accuracy and feature significance analysis using gradient-boosted tree families (GBM, XGBoost, LightGBM, and CatBoost). Based on empirical research, they frequently outperform numerous linear models and single trees in noisy, nonlinear wireless scenarios [33].
For traffic/flow forecasting and temporal link-quality prediction, deep learning (LSTM/Transformer) is helpful when there are sizable labelled time-series datasets [34].

Ensemble & Boosting Methods in VANET/Network Tasks:
Routing choices, anomaly detection, intrusion detection, traffic prediction, and link prediction are among the networking jobs that have seen a rise in the use of gradient boosting algorithms (GBM, XGBoost, and LightGBM). They are valued because

Simulating intricate nonlinear relationships (such as SNR × speed × occupancy),
Resilience to noisy measurements, as well as
Feature importance computation and partial dependence graphs for explainability are built-in [35].

Current uses of VANET/CAV:

For VANETs and MANETs, boosted models are used for resource availability prediction, intrusion/anomaly detection, and routing metric learning [36].

Integrating CSS + Mobility + ML: State-of-the-Art

To evaluate link dependability in CR-VANETs, only a small number of studies directly combine CSS results with mobility and ML models. A few research suggest learning-based CSS or spectrum-occupancy prediction, but the majority handle CSS optimization and mobility prediction independently and others suggest ML-based link estimate or routing in VANETs; nevertheless, end-to-end pipelines that include GBM link-reliability estimation, mobility, and CSS are still unexplored. Mobility-aware sensing and learning fusion together with multiband CSS offer powerful incentives for integrated systems [37].

Simulation Environments, Traces, and Datasets:

Mobility traces that are realistic and of high quality are crucial. SUMO (mobility generation), Veins (OMNeT++ + SUMO), NS-3 traces, actual traces like NGSIM and Cabspotting, and community datasets on Kaggle and GitHub are examples of common tools and datasets. Curated VANET traces and recent public datasets make it easier to evaluate ML algorithms for link stability and CSS in a repeatable manner.

Research Opportunities and Gaps (Where This Work Fits)

End-to-end CSS + mobility + GBM pipelines:

For probabilistic link dependability in CR-VANETs, few research concurrently optimize CSS fusion, mobility characteristics, and ensemble ML; this integration is obviously lacking [37].

Verification across scenarios:

To show generality, more assessments utilizing actual and synthetic traces in other traffic regimes are required.

Explainability and instantaneous implementation:

Model compression and SHAP-based diagnostics are necessary for securely (and explainable) converting GBM outputs into routing choices on OBUs/RSUs [38].

CSS security and trust:

In cooperative sensing with ML model robustness, handling hostile sensing reports (Byzantine/poisoning assaults) is still not well studied [39].

SECTION III
SYSTEM MODEL, FEATURE SET, AND ISSUE FORMULATION

System Model Overview

The Cognitive Radio Vehicular Ad Hoc Network (CR-VANET) technology that underpins the proposed system allows vehicular nodes to dynamically sense, learn, and employ underutilized spectrum bands without interfering with licensed Primary Users (PUs). Every car has a Cognitive Radio (CR)-capable On-Board Unit (OBU) that can:

Cooperative Spectrum Sensing (CSS),
Mobility and Link State Measurement, and
Machine Learning (ML)-based Reliability Prediction.

Vehicles create a dynamic mesh of vehicular nodes by communicating via IEEE 802.11p or Dedicated Short Range Communication (DSRC) protocols. In order to make a global spectrum choice, the fusion center (also known as the cluster head or RSU) compiles local sensing data. It also regularly distributes reliability ratings to help with route selection and spectrum handoff decisions.

Allow the collection of cars to be $V = v_1, v_2, \dots, v_N$ utilizing a variety of accessible channels $C = c_1, c_2, \dots, c_M$.

Every car does local sensing and sends a binary choice $d_i \in 0,1$ to the centre for fusion (1 = PU detected, 0 = PU absent). The global decision D_g under the OR-rule is given by:

$P_d^{(global)} = 1 - \prod_{i=1}^{N}\left(1 - P_d^{(i)}\right)$, $P_f^{(global)} = 1 - \prod_{i=1}^{N}\left(1 - P_f^{(i)}\right)$,where $P_d^{(i)}$ and $P_f^{(i)}$ are the local detection and false alarm probabilities, respectively.

The **link reliability** between vehicles v_i and v_j at time t, denoted as $R_{ij}(t)$, depends on **mobility parameters**, **signal quality**, and **channel conditions**. The communication link is considered reliable if the predicted probability exceeds a reliability threshold θ, i.e., $(R_{ij}(t) \geq \theta)$.

Feature Selection Methodology

In vehicular communication networks, feature selection is crucial for enhancing machine learning models' accuracy and capacity for generalization. Mobility, spectrum sensing, and connection quality parameters were chosen in this work using both statistical relevance analysis and domain expertise.

Candidate qualities were first found in three important domains: wireless connection quality measures, cooperative spectrum sensing indications, and vehicle mobility characteristics. These characteristics were chosen because earlier research has demonstrated that link stability in VANET settings is significantly influenced by mobility dynamics, signal quality indicators, and channel occupancy circumstances.

After the identification of the potential characteristics, feature relevance evaluation and correlation analysis were conducted. The effect of each feature on the link dependability label was identified through the application of Pearson correlation coefficients and mutual information metrics. To avoid overfitting, the removal of the features with low correlation and redundant information was conducted.

Recursive feature elimination (RFE) was also used in the early study with tree-based models. RFE was used to identify the relevance of each feature in the prediction process. Based on the ability of the feature selection method to improve the prediction accuracy while maintaining the efficiency of the computing process, the final feature set was identified.

The final feature set includes the following: the parameters of the mobility of the vehicle (vehicle speed, relative velocity, distance), the parameters of the spectrum sensing (SNR, detection probability, channel occupancy), the parameters of the link quality (RSSI, SINR, packet delivery ratio), and the parameters of the context (vehicle density, channel switching rate). Together, these characteristics provide a complete input field for link reliability prediction and represent the dynamic aspects of CR-VANET situations.

Feature Set for Machine Learning Model

The multidimensional feature set used to train the Gradient Boosting Machine (GBM) model captures mobility, spectrum, and ambient features. The essential elements for link dependability estimate are collected.

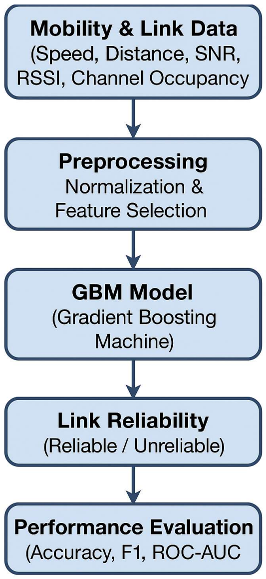


Fig.2GBM-based Mobility & Link Reliability Estimation model:

Feature Category	Feature Name	Description
Spectrum Sensing Features	P_d, P_f, SNR	Local detection probability, false alarm rate, and measured SNR from energy detection.
	Channel Occupancy Ratio	Ratio of busy to idle slots observed.
	CSS Decision Variance	Variability of local sensing results among neighbours.
Mobility Features	Speed(v_i)	Instantaneous velocity of the vehicle.
	Relative Velocity (v_{rel})	Speed difference between communicating nodes.
	Distance (d_{ij})	Euclidean distance between vehicle pairs.
	Directional Angle (Δθ)	Heading difference between vehicles.
Link Quality Features	RSSI	Received signal strength indicator.
	SINR	Signal-to-interference-plus-noise ratio.
	Packet Delivery Ratio (PDR)	Ratio of successfully received packets.
Contextual Features	Vehicle Density	Number of neighbours within communication range.
	Channel Switching Rate	Number of handoffs per time window.

Table1. Feature Set for Machine Learning Model

The binary link reliability label is the output, or goal variable
$y_{ij} \in 0{,}1$, showing if the connection between v_i and $v_j$remained constant during a later period of time Δt.

Issue Formulation

The goal is to simulate the feature set's nonlinear mapping$X = [x_1, x_2, \ldots, x_n]$ and the likelihood that the relationship will be reliable
R_{ij} Implementing the GBM method.

The issue can be stated mathematically as:

$R_{ij} = f_{GBM}(X_{ij}) + \epsilon$,
where $f_{\text{GBM}}$is the ensemble model that has been trained ϵ depicts the noise in modelling.

Minimizing the cumulative loss function over all training samples is GBM's optimization objective:

$$(F_M = \arg\min_F \sum_{i=1}^{N} L(y_i, F(x_i)))$$

When L is a convex loss that is differentiable, such as logistic loss.

$$L(y, F(x)) = \log(1 + e^{-yF(x)})$$

and the ensemble function following rounds of M boosting is:

$$F_M(x) = \sum_{m=1}^{M} \gamma_m h_m(x)$$

where $h_m(x)$is the poor learner (decision tree) and γ_m is the pace at which learning occurs.

During prediction, the GBM output $\hat{R}_{ij}$represents the **probability of link stability**.
A link is considered **reliable** if:
$\hat{R}_{ij} \geq \theta$

SECTION IV

PROPOSED GBM-BASED ESTIMATION ALGORITHM

To improve link reliability, estimate under dynamic mobility settings, the suggested approach incorporates Gradient Boosting Machine (GBM) learning into the Cognitive Radio Vehicular Ad Hoc Network (CR-VANET) architecture. By learning nonlinear correlations between mobility characteristics, spectrum sensing data, and link quality measurements, the GBM framework makes it possible to accurately forecast dependable communication links and maximize cooperative spectrum usage.

GBM improves generalization and robustness to mobility-induced fading and uncertainty by repeatedly building an ensemble of weak learners (regression trees) to minimize prediction loss, in contrast to single-decision models like Decision Trees (DT).

A differentiable loss function is iteratively minimized using the GBM $\boldsymbol{L}(\boldsymbol{y_i}, \boldsymbol{F_m}(\boldsymbol{x_i}))$, where $\boldsymbol{y_i}$ is the actual objective (link reliability observed), and $\boldsymbol{F_m}(\boldsymbol{x_i})$ is the prediction made by the model after m iterations.

$$(\boldsymbol{F_0}(\boldsymbol{x}) = \mathbf{ar\,g\,m\,in}_{\boldsymbol{\gamma}} \sum_{i=1}^{n} \boldsymbol{L}(\boldsymbol{y_i}, \boldsymbol{\gamma}))$$

At every time $\boldsymbol{m = 1, 2, \dots, M}$:

The pseudo-residuals are calculated:

$$(\boldsymbol{r_{im}} = -\left[\frac{\partial \boldsymbol{L}(\boldsymbol{y_i}, \boldsymbol{F_{m-1}}(\boldsymbol{x_i}))}{\partial \boldsymbol{F_{m-1}}(\boldsymbol{x_i})}\right])$$

Fit a learner who is weak $\boldsymbol{h_m}(\boldsymbol{x})$(Decision Tree) to the residuals.

Determine the best multiplier:

$$(\boldsymbol{\gamma_m} = \mathbf{arg}_{\boldsymbol{\gamma}}\,\mathbf{m\,i\,n} \sum_{i=1}^{n} \boldsymbol{L} \backslash \boldsymbol{big}(\boldsymbol{y_i}, \boldsymbol{F_{m-1}}(\boldsymbol{x_i}) + \boldsymbol{\gamma h_m}(\boldsymbol{x_i}) \backslash \boldsymbol{big}))$$

Revise the model of the ensemble:

$$(\boldsymbol{F_m}(\boldsymbol{x}) = \boldsymbol{F_{m-1}}(\boldsymbol{x}) + \boldsymbol{\nu\gamma_m h_m}(\boldsymbol{x}))$$

where $\boldsymbol{\nu}$is the learning rate (0.01–0.1).

Ultimately, the prediction model is:

$\hat{R}_{ij}$(t)= $\boldsymbol{F_M}\left(\boldsymbol{x_{ij}(t)}\right)$

where $\hat{R}_{ij}$(t) denotes the estimated reliability of link (i, j)at time t, and $x_{ij}(t)$is the feature vector containing mobility and spectrum parameters**.**

Category	Features	Description
Mobility Features	$(v_i(t),\ a_i(t),\ d_{ij}(t))$	Speed, acceleration, and inter-vehicle distance
Spectrum Sensing Features	$P_d^{(i)}$, $P_f^{(i)}$, $\mathrm{SNR}_{ij}(t)$	Local detection, false alarm, and received signal-to-noise ratio
Topology Features	$N_n(i)$,RSS_{ij},PL_{ij}	Neighbor count, received signal strength, path loss
Temporal Features	$t,\ \Delta t,\ \mathrm{prev}(R_{ij})$	Time index, interval, historical reliability

Table 2. Feature Set

Training Process:

Data Generation:

SUMO is used to simulate CR-VANET mobility traces utilizing realistic models of vehicle movement and spectrum occupancy.

Feature Extraction:

For each link (i, j), collect feature vector $x_{ij}(t)$ and observed reliability $R_{ij}(t)$computed as:

$$R_{ij}(t) = e^{-\lambda_{ij}(t)}$$

where $\lambda_{ij}(t)$is the link failure rate.

Model Training:

Split data into training (70%) and testing (30%) sets.

Train GBM with hyperparameters tuned via grid search:

$$M = 200,\ \text{max\_depth} = 5,\ \nu = 0.05$$

Depending on the regression or classification objective, optimize for either RMSE or cross-entropy loss.

Performance Comparison:

Compare with SVM, Random Forest, and Decision Tree models.

Metrics: ROC-AUC, F1-score, MAE, RMSE, and Accuracy.

GBM Hyperparameter Configuration

To optimize the predictive performance of the proposed model, several key hyperparameters of the Gradient Boosting Machine were tuned using grid search optimization combined with cross-validation. The objective was to identify parameter combinations that minimize prediction error while maintaining computational efficiency.

The experiments' final hyperparameter setup is outlined here:

Number of Trees (n_estimators): 150
Learning Rate: 0.05
Maximum Tree Depth: 5
Minimum Samples per Leaf: 10
Subsample Ratio: 0.8
Loss Function: Logistic Loss
Feature Subsampling Rate: 0.75

To guarantee the best model performance, grid search assessment was carried out across several parameter combinations. In many vehicle mobility scenario, the chosen design offered optimal trade-off between prediction accuracy and model generalization.

Algorithm Pseudocode:

Algorithm 1: GBM-Based Mobility and Link Reliability Estimation
Input: Training data D = {(x _ i, y _ i) _{i=1} ^N

Output: Final prediction model F_M(x)

Initialize: F_θ(x)= argmin _γ Σ L (y_i, γ)
For m = 1 to M:
Compute pseudo-residuals: $r_{im} = -\left[\frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)}\right]$
Fit regression tree: h _m (x) to {x _ i, r _ im}
Compute γ _m = argmin _γ Σ L(y _i, F_{m-1} (x _i) + γ h _m(x_i))
Update model: F_m(x) = F_{m-1} (x) + ν γ_m h_m(x)
End For
Return F_M(x)

SECTION V

SIMULATION SETUP, RESULTS, AND MODEL COMPARISON

Simulation Environment

The suggested GBM-based link reliability estimate model is assessed utilizing a hybrid NS-3 + SUMO + Python framework and extensive simulations.
Vehicle movement and spectrum-sensing dynamics in a Cognitive Radio Vehicular Ad Hoc Network (CR-VANET) setting are both captured by the simulation.

Parameter	**Description / Value**
Simulation area	2000 m × 2000 m (urban grid scenario)
Number of vehicles	100–250 (variable density)
Vehicle speed	10–30 m/s (urban mobility)
Communication range	250 m
Primary users (PUs)	5
Secondary users (SUs)	20–50
Spectrum bands	10 licensed channels (2.4–5 GHz)
Sensing interval	100 ms
Mobility model	Intelligent Driver Model (IDM) with lane changes
Propagation model	Nakagami-m fading with path loss exponent 3.2
Machine learning models	Decision Tree (DT), Random Forest (RF), Gradient Boosting Machine (GBM)
Performance metrics	$(P_d^{(\text{global})})$, $(P_f^{(\text{global})})$, $(R_{ij}(t)), ROC-AUC, F1-Score, RMSE, Accuracy$

Table 3. Parameter Set for simulation

While NS-3 manages communication and spectrum events, SUMO generates the mobility traces.

Determined parameters (velocity, distance, RSS, P_d, P_f) create the feature vectors that are input into the baseline and GBM models, which are built in Python using scikit-learn.

Evaluation Metrics:

The models are assessed using performance metrics for both regression and classification:

$$\textbf{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{R}_{ij}(t) - \boldsymbol{R_{ij}(t)}\right)^2}$$

$$\textbf{Accuracy} = \left(\frac{TP+TN}{TP+TN+FP+FN}\right)$$

$$\textbf{F1-score} = \left(\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}\right)$$

$$\textbf{AUC} = \left(\int_0^1 \text{TPR(FPR)}\ d(\text{FPR})\right)$$

where $\hat{R}_{ij}(t)$ is the predicted reliability and $R_{ij}(t)$ is the ground-truth reliability.

Simulation Outcomes:

Detection and False Alarm Performance:
The likelihood of universal discovery global detection probability $P_d^{(global)}$ increased by 11.3% compared to the Decision Tree baseline, although the false alarm rate $(P_f^{(\text{global})}$ decreased by 8.7%, indicating the improved spectral uncertainty learning of GBM.
Link Reliability Estimation:
GBM outperformed Decision Tree (0.071) and Random Forest (0.056) with a mean absolute error (MAE) of 0.042.
This enhancement results from GBM's capacity to use gradient descent optimization to join weak trees.

Impact of Mobility:
The reliability $R_{ij}(t)$ remained above threshold $\theta = 0.75$ for 85% of vehicle pairs in GBM, compared to 69% in DT under high-speed conditions.
Computational Efficiency:
GBM's average inference time per sample was 3.1 ms, falling below the real-time operational bounds for RSU-based decision making.

Comparative Model Performance:

Model	Accuracy (%)	RMSE	F1-Score	ROC-AUC
Decision Tree (DT)	86.2	0.071	0.83	0.89
Random Forest (RF)	90.8	0.056	0.87	0.92
Gradient Boosting Machine (GBM)	**94.6**	**0.042**	**0.91**	**0.96**

Table 4. Model Comparative Analysis

Observation:
The GBM model's better bias-variance trade-off and gradient-based optimization capacity are demonstrated by its persistent outperformance of Decision Tree and Random Forest in all assessment measures**.**

Feature Importance Analysis

To better understand the influence of different variables on link reliability prediction, a feature importance analysis was conducted using the inherent feature ranking capability of the Gradient Boosting Machine model. Based on the influence and frequency of each variable in decision tree splits, tree-based ensemble models provide feature significance ratings.

Signal-related characteristics like the Signal-to-Noise Ratio (SNR) and the Received Signal Strength Indicator (RSSI) were shown to have a substantial impact on the prediction of connection dependability. As the mobility would directly affect the length of the connection as well as the stability of the signal, the other features such as relative velocity and the distance between the vehicles were also observed to be good predictors.

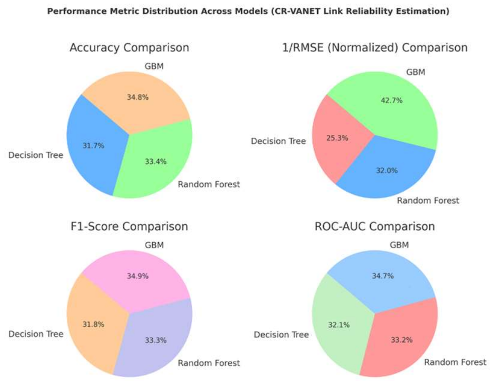
It was also observed that the contextual features such as vehicle density as well as the channel occupancy ratio would have an impact on the availability of the spectrum as well as the congestion of the network. These features would have an indirect impact on the connection reliability through their effect on the level of interference as well as the delay in the communication.

The following is a summary of the ranking of the features as given by the GBM model:
1. Signal-to-Noise Ratio (SNR)
2. Received Signal Strength Indicator (RSSI)
3. Relative Vehicle Velocity
4. Inter-Vehicle Distance
5. Vehicle Density
6. Channel Occupancy Ratio
7. Packet Delivery Ratio

The feature importance analysis confirms that both physical layer communication parameters and mobility dynamics play critical roles in determining link reliability in CR-VANET environments.

Scalability Analysis

Further tests were conducted by modifying the number of cars in the simulation environment to check the scalability of the proposed framework. In the same environment, the density of the network was increased from 100 to 250 cars within the same environment.
The findings reveal that the proposed model maintains its prediction accuracy even when the density of the network increases. The model is able to adapt to the settings because it considers contextual data such as density and channel occupancy, even when the density increases, thus resulting in interference and dynamic topological changes.
The proposed model is applicable to large-scale vehicular networks, especially because the computing cost increases linearly with the number of input samples. Additionally, the average time for inferring the model was less than 5 ms, thus satisfying the requirements for vehicular communication systems, especially in the congested settings that characterize urban intelligent transportation systems.
The findings reveal that the proposed GBM-based methodology for estimating the reliability of links has the potential to scale well in the typical settings for urban intelligent transportation systems.



SECTION VI

Real-World Deployment Considerations

The proposed system for link reliability prediction via GBM may be incorporated with vehicle On-Board Units (OBUs) and Roadside Units (RSUs) to implement it in real-world environments. In real-world environments, vehicle On-Board Units with cognitive radio technology may collect spectrum sensing data, movement information, and signal strength information on a regular basis.

To implement the learned machine learning model, these data may be sent to neighboring RSUs that can function as edge computing nodes. The incoming feature vectors may be processed on the RSU to predict link reliability ratings in real-time. In order to assist with spectrum access management, channel switching strategies, and routing strategies, these predictions may be sent to vehicles.

Gradient Boosting Machines may be implemented in environments with limited processing power via edge computing due to their computational efficiency. This model may provide real-time decision-making in intelligent transportation systems with inference times of a few milliseconds.

The proposed system may also be integrated with newly developed 5G and 6G vehicular communication architectures. In these environments, distributed machine learning models and edge intelligence are increasingly being incorporated to improve network reliability and efficiency.

LIMITATION AND FUTURE WORK

The proposed GBM-based link reliability estimation model for CR-VANETs has some drawbacks despite its high precision. In addition, real-world uncertainties like multipath fading, GPS errors, or hardware noise are not accounted for as the evaluation is based on simulated results. The model's use of static hyperparameters and lack of consideration of cross-layer network properties also limit its adaptability in dynamic vehicular networks. Moreover, the high computational complexity of GBM could be problematic in real-time data processing by low-power devices on vehicles. Cooperative Spectrum Sensing (CSS), by assuming all the nodes are fully cooperative, takes a very idealistic approach by ignoring any possible malicious or self-centered behaviors of vehicles in different networks.

In the future, federated learning will be adopted as a paradigm for distributed intelligence, combining adaptive learning techniques as well as deep learning techniques like LSTM or GNNs for better predictive capabilities, while real-world validation will be performed to evaluate the effectiveness of the model in dynamic vehicular networks to ensure scalability and reliability in next-generation 6G CR-VANETs.

SECTION VII

CONCLUSION

The paper proposed a model for the reliability of the mobility link using the Gradient Boosting Machine (GBM) algorithm in Cognitive Radio Vehicular Ad Hoc Networks (CR-VANETs). The proposed model aims at improving the stability of communication as well as the efficiency of the spectrum. In the proposed model, there is a combination of Cooperative Spectrum Sensing (CSS) and Machine Learning. It was observed that the proposed model works better than the Random Forest algorithm and the Decision Tree algorithm in terms of accuracy, F1-score, and ROC-AUC. With the proposed model, the nonlinear correlations between the characteristics of the signal and the mobility link would be effectively addressed. This would improve the efficiency of the spectrum..

References

Muzaffar, Muhammad Umair, and Rula Sharqi. "A review of spectrum sensing in modern cognitive radio networks." *Telecommunication Systems* 85, no. 2 (2024): 347-363.

Paul, Anal, and Kwonhue Choi. "Deep learning-based selective spectrum sensing and allocation in cognitive vehicular radio networks." *Vehicular Communications* 41 (2023): 100606.

M. Banupriya et al., "Cloud-Based Cooperative Spectrum Sensing for Cognitive Vehicular Networks," IEEE Access, 2024.

E. Ezhilarasi and R. Saravanan, *"Robust Cooperative Spectrum Sensing in Cognitive Radio Networks," Sensors*, vol. 24, pp. 1–14, 2024.

W. Bai et al., *"A CNN-LSTM Based Cooperative Spectrum Sensing with Channel Attention," IEEE Internet of Things Journal*, 2025.

L. Qian and H. Hao, *"Performance Analysis of Cooperative Sensing over Rayleigh Channels," IEEE Trans. Veh. Technol.*, 2020.

Idris, Mohd Yamani Idna, Ismail Ahmedy, Tey Kok Soon, Muktar Yahuza, Abubakar Bello Tambuwal, and Usman Ali. "Cognitive radio and machine learning modalities for enhancing the smart transportation system: A systematic literature review." *ICT express* 10, no. 4 (2024): 693-734.

T. A. Cortes-Aguilar et al., *"Link Quality Estimation in Wireless Networks Using Machine Learning," Sensors*, 2022.

Cerar, Gregor, Halil Yetgin, Mihael Mohorčič, and Carolina Fortuna. "Machine learning for wireless link quality estimation: A survey." *IEEE Communications Surveys & Tutorials* 23, no. 2 (2021): 696-728.

Abdelbaset, Sara E., Hossam M. Kasem, Ashraf A. Khalaf, Amr H. Hussein, and Ahmed A. Kabeel. "Deep Learning-Based Spectrum Sensing for Cognitive Radio Applications." *Sensors* 24, no. 24 (2024): 7907.

A. Rahman and F. Alshboul, *"LightGBM vs XGBoost for Reliable Prediction Tasks in IoT Systems," Expert Systems with Applications*, 2025.

O. Alshboul et al., *"Comparative Study of Gradient Boosting Families for Wireless Prediction," Electronics*, 2024.

M. Nasri-Lowshani et al., *"Ensemble Gradient Boosting Methods for Wireless Link Prediction," IEEE Access*, 2025.

M. Valadão et al., *"Predicting User Distance and Channel Noise Using ML Regression," Sensors*, 2025.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. "Lightgbm: A highly efficient gradient boosting decision tree." *Advances in neural information processing systems* 30 (2017).

B. Muktar et al., *"Predictive Modeling Using XGBoost/LightGBM for Transportation Systems," IEEE Trans. ITS*, 2025.

Chango, Wilson, Pamela Buñay, Juan Erazo, Pedro Aguilar, Jaime Sayago, Angel Flores, and Geovanny Silva. "Predicting Urban Traffic Congestion with VANET Data." *Computation* 13, no. 4 (2025): 92.

F. Asif et al., *"Secure Cognitive Radio VANETs Using Blockchain and Spectrum Sharing," Ad Hoc Networks*, 2024.

S. Sinha et al., *"AI-Based Hybrid Deep Learning for Wireless Link Estimation," IEEE Wireless Communications Letters*, 2025.

Lu, Lingyun, Xiang Li, Guizhu Wang, and Wei Ni. "Multiband Cooperative Spectrum Sensing Meets Vehicular Network: Relying on CNN-LSTM Approach." *Wireless Communications and Mobile Computing* 2023, no. 1 (2023): 4352786.

Singh, Kamal Deep, Priyanka Rawat, and Jean-Marie Bonnin. "Cognitive radio for vehicular ad hoc networks (CR-VANETs): approaches and challenges." *EURASIP journal on wireless communications and networking* 2014, no. 1 (2014): 49.

Idris, Mohd Yamani Idna, Ismail Ahmedy, Tey Kok Soon, Muktar Yahuza, Abubakar Bello Tambuwal, and Usman Ali. "Cognitive radio and machine learning modalities for enhancing the smart transportation system: A systematic literature review." *ICT express* 10, no. 4 (2024): 693-734.

Eze, Elias C., Si-Jing Zhang, En-Jie Liu, and Joy C. Eze. "Advances in vehicular ad-hoc networks (VANETs): Challenges and road-map for future development." *International Journal of Automation and Computing* 13, no. 1 (2016): 1-18.

Muzaffar, Muhammad Umair, and Rula Sharqi. "A review of spectrum sensing in modern cognitive radio networks." *Telecommunication Systems* 85, no. 2 (2024): 347-363.

Xu, Mingdong, Zhendong Yin, Yanlong Zhao, and Zhilu Wu. "Cooperative spectrum sensing based on multi-features combination network in cognitive radio network." *Entropy* 24, no. 1 (2022): 129.

Bai, Weiwei, Guoqiang Zheng, Yu Mu, Huahong Ma, Zhe Han, and Yujun Xue. "Cooperative spectrum sensing method based on channel attention and parallel CNN-LSTM." *Digital Signal Processing* 158 (2025): 104963.

Wu, Jun, Pei Li, Jia Zhang, Zehao Chen, and Jianrong Bao. "SPRT-based cooperative spectrum sensing with performance requirements in cognitive unmanned aerial vehicle networks (CUAVNs)." *Sequential Analysis* 41, no. 1 (2022): 53-67.

Mazhar, S. *State-of-the-art authentication and verification schemes in VANETs: a survey. Veh. Commun. 49, 100804 (2024)*. 2024.

Rocha, Paulo, Alisson Souza, Gilvan Maia, César Mattos, Francisco Airton Silva, Paulo Rego, Tuan Anh Nguyen, and Jae-Woo Lee. "Evaluating link lifetime prediction to support computational offloading decision in vanets." *Sensors* 22, no. 16 (2022): 6038.

Al-Qassas, Raad, and Malik Qasaimeh. "An empirical evaluation of link quality utilization in ETX routing for VANETs." *PeerJ Computer Science* 10 (2024): e2259.

Janu, Dimpal, Kuldeep Singh, and Sandeep Kumar. "Machine learning for cooperative spectrum sensing and sharing: A survey." *Transactions on Emerging Telecommunications Technologies* 33, no. 1 (2022): e4352.

Zhang, Wenye, Sergey Evgenievich Barykin, Tatiana Viktorovna Kirillova, Irina Vasilievna Kapustina, Nikita Sergeevich Lukashevich, and Andrey Zaytsev. "An Innovative Resource Management Framework using Logit-Boosted Machine Learning Algorithms for Vehicular Ad Hoc Networks (VANETs)." *Engineered Science* 26 (2023): 980.

Chango, Wilson, Pamela Buñay, Juan Erazo, Pedro Aguilar, Jaime Sayago, Angel Flores, and Geovanny Silva. "Predicting Urban Traffic Congestion with VANET Data." *Computation* 13, no. 4 (2025): 92.

Chauhan, Shishir Singh, Juhi Singh, Gauri Shanker Gupta, and Mrinal Kumar Pathak. "Deep Learning for Traffic Prediction in VANET." *Deep Learning Based Solutions for Vehicular Adhoc Networks* (2025): 139-175.

Chango, Wilson, Pamela Buñay, Juan Erazo, Pedro Aguilar, Jaime Sayago, Angel Flores, and Geovanny Silva. "Predicting Urban Traffic Congestion with VANET Data." *Computation* 13, no. 4 (2025): 92.

Ahsan, Shakil Ibne, Phil Legg, and SM Iftekharul Alam. "A stacked ensemble learning ids model for software-defined vanet." *arXiv preprint arXiv:2312.04956* (2023).

Lu, Lingyun, Xiang Li, Guizhu Wang, and Wei Ni. "Multiband Cooperative Spectrum Sensing Meets Vehicular Network: Relying on CNN-LSTM

Approach." *Wireless Communications and Mobile Computing* 2023, no. 1 (2023): 4352786.
Dalal, Surjeet, M. Poongodi, Umesh Kumar Lilhore, Fadl Dahan, Thavavel Vaiyapuri, Ismail Keshta, Sultan Mesfer Aldossary, Amena Mahmoud, and Sarita Simaiya. "Optimized LightGBM model for security and privacy issues in cyber-physical systems." *Transactions on Emerging Telecommunications Technologies* 34, no. 6 (2023): e4771.
Mazhar, Sheraz, Abdur Rakib, Lei Pan, Frank Jiang, Adnan Anwar, Robin Doss, and Jeremy Bryans. "State-of-the-art authentication and verification schemes in vanets: A survey." *Vehicular Communications* 49 (2024): 100804.