

Trust-Aware Kalman Filtering For Secure Traffic Prediction From Social Media Streams

Ahmad Faisal Abidin¹, Nor Surayati Mohamad Usop², Mohd Fadzil Abdul Kadir³,
Mohamad Afendee Mohamed⁴

^{1 2 3 4} Department of Computer Science, Faculty of Informatics and Computing, Universiti Sultan
Zainal Abidin, 22200 Besut, Terengganu, Malaysia

¹ <https://orcid.org/0000-0003-0186-5295>; ² <https://orcid.org/0009-0007-8133-1416>; ³
<https://orcid.org/0000-0002-3647-5134>;

⁴ <https://orcid.org/0000-0001-5985-3970>

*Corresponding Author: Ahmad Faisal Abidin

Abstract— The growing use of social media platforms, particularly Twitter, for real-time reporting of road traffic conditions presents new opportunities for intelligent transportation systems. However, traffic data derived from Twitter is inherently noisy, unstructured, and susceptible to misinformation, which poses significant challenges to prediction accuracy and data security. This review paper provides an analytical examination of Kalman Filter (KF)-based models for enhancing the reliability, robustness, and integrity of traffic information extracted from Twitter. The study focuses on variations of the Kalman Filter, with particular emphasis on the Extended Kalman Filter (EKF), and evaluates their effectiveness in handling uncertain, incomplete, and dynamically evolving data patterns typical of social media streams. Furthermore, the review discusses the role of KF-based prediction in supporting anomaly detection mechanisms for identifying malicious, manipulated, or bot-generated traffic reports, thereby improving data authenticity and trustworthiness. Existing studies indicate that Kalman Filtering techniques enhance traffic flow prediction accuracy, reduce noise in user-generated content, and improve model stability in rapidly changing traffic environments. Despite these advantages, several challenges remain, including scalability, semantic interpretation of textual data, and integration with cybersecurity frameworks such as sentiment analysis, machine learning-based classification, and blockchain-enabled validation. The paper concludes that Kalman Filter models hold strong potential for securing and predicting Twitter-sourced traffic data; however, further interdisciplinary research is required to develop optimized hybrid systems that integrate natural language processing, advanced filtering techniques, and robust security frameworks. The increasing reliance on social media platforms such as Twitter for real-time traffic monitoring presents both opportunities and challenges for intelligent transportation systems (ITS). While Twitter enables rapid dissemination of traffic-related information, the data is inherently noisy, unstructured, and vulnerable to misinformation and malicious manipulation. This paper proposes a secure Kalman Filter-based framework for trust-aware traffic prediction using Twitter-derived data. The proposed approach integrates simulated traffic data from SUMO with real-time social media observations through a data fusion mechanism governed by the Kalman Filter. The Extended Kalman Filter (EKF) is employed to model nonlinear traffic dynamics, while trust-aware filtering is achieved through adaptive Kalman Gain adjustment. Security mechanisms including digital signatures, blockchain validation, and anomaly detection are incorporated to enhance data integrity. Experimental results demonstrate improved prediction accuracy and robustness against unreliable inputs.

Keywords— Kalman Filtering, Traffic Prediction, Data Security, Social Media Analytics, Twitter Traffic Data.

I. INTRODUCTION

Urban traffic management has increasingly relied on data-driven approaches to address congestion, enhance mobility, and improve road safety. Traditional sensing infrastructures such as inductive loops, cameras, and radar systems, although widely deployed, remain limited by cost, scalability, and vulnerability to environmental disturbances [1]. In recent years, social media platforms particularly Twitter have emerged as valuable supplementary sources of real-time traffic information, offering

rapid crowdsourced updates from road users themselves [2]. This shift toward user-generated data has contributed to a new class of intelligent transportation systems (ITS) where conventional sensor data and social data streams converge to enhance situational awareness [3].

However, Twitter-derived traffic data is often unstructured, noisy, and affected by linguistic ambiguity, making accurate prediction and reliable interpretation challenging [4]. To address these issues, researchers have increasingly turned to Kalman Filter (KF) models due to their mathematical robustness and ability to estimate system states in the presence of uncertainty [5][6]. The Kalman Filter and its variants such as the Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and Ensemble Kalman Filter (EnKF) have been widely applied in transportation modelling for tasks including traffic flow estimation, travel-time prediction, incident detection, and anomaly recognition [5]. Their recursive nature makes them well-suited for streaming environments where continuous updates are required, a characteristic that aligns with the dynamic nature of Twitter data [2].

Despite the significant interest in applying Kalman-based models to social media derived traffic information, challenges remain regarding data quality, reliability, and security [7]. Social media platforms are vulnerable to misinformation, data manipulation, bot-generated content, and coordinated malicious activities [8]. These issues can distort traffic prediction models and compromise transportation management decisions if not properly addressed. Moreover, the integration of crowdsourced data into critical ITS ecosystems raise concerns regarding data authenticity, integrity, and resilience against adversarial interference [9][10]. As transportation systems become more connected and dependent on digital sources, cybersecurity considerations become an essential component of traffic data analytics.

Given these complexities, a comprehensive analytical review of Kalman Filter models evaluating both their predictive performance and security implications when applied to Twitter-derived traffic data is timely and necessary. While previous studies have examined the use of Kalman Filters in traffic estimation [11][12] and others have explored the utility of social media for traffic monitoring [13], there is limited consolidated work focusing specifically on the intersection of prediction modelling and security concerns within this context. Additionally, the rapid growth of social sensing technologies has produced a diverse set of methodological approaches, yet no clear synthesis exists to evaluate their comparative strengths, limitations, and suitability for secure ITS environments [14].

This review addresses these gaps by performing an analytical examination of existing Kalman Filter models applied to Twitter-based traffic prediction. The analysis includes classical KF approaches, nonlinear extensions, and hybrid methods that combine Kalman Filters with machine learning algorithms, natural language processing (NLP) techniques, and anomaly detection frameworks. These hybrid models are particularly relevant as they attempt to leverage the strengths of statistical filtering and data-driven learning to handle the complexity of social media data streams [15]. Furthermore, the review investigates how Kalman Filters can contribute not only to prediction accuracy but also to security-related tasks such as detecting abnormal patterns, verifying data consistency, and mitigating the influence of malicious or misleading posts.

Overall, the increasing availability of user-generated traffic information presents an opportunity to enhance existing ITS frameworks, but only if challenges related to noise, uncertainty, and security are properly addressed. Kalman Filter models, known for their robustness in uncertain environments, hold significant promise in this area yet require

careful analysis and adaptation when deployed on social media data. This paper contributes to the body of knowledge by synthesizing current research, identifying major gaps, and proposing future research directions for secure and reliable traffic prediction using Kalman-based approaches. The findings are expected to benefit researchers, practitioners, and policymakers seeking to develop next-generation intelligent transportation systems that are both data-rich and security-conscious.

Research Gap and Contributions:

This paper addresses the lack of integration between prediction accuracy, noise filtering, and data security in social media-based traffic systems.

Contributions:

1. Kalman Filter-based fusion of SUMO and Twitter data.
2. Trust-aware filtering using adaptive Kalman Gain.
3. Integration of blockchain and digital signature validation.
4. Experimental validation of robustness and accuracy.

II. RELATED WORKS

With the increased reliance on smart mobility systems and data-driven transportation planning, social media platforms have emerged as valuable supplementary sources of real-time traffic information. Among them, Twitter has become a widely used medium where users share traffic incidents, congestion levels, delays, and emergencies. Compared to traditional sources such as loop detectors, surveillance cameras, or GPS vehicle traces, Twitter offers faster reporting latency and broader geographic coverage. However, Twitter-generated traffic data is inherently noisy, unstructured, and inconsistent due to its textual nature, human subjectivity, and susceptibility to misinformation and automated bot-generated content. These characteristics introduce challenges in maintaining data credibility, prediction accuracy, and system dependability in traffic analytics applications [16].

To address such uncertainties, Kalman Filter (KF) models have been applied across various domains requiring real-time estimation under noise and incomplete data. Originally developed for aerospace navigation, KF algorithms have demonstrated exceptional capabilities in state estimation under dynamic environmental conditions [17][18]. In this research, the Extended Kalman Filter (EKF) is selected due to its suitability for modeling nonlinear traffic dynamics while maintaining computational efficiency for real-time applications [19]. Traffic conditions inferred from Twitter data exhibit inherent nonlinearities arising from irregular posting behavior, varying traffic flow patterns, and noisy, incomplete observations. Unlike the standard Kalman Filter, which assumes linear system and measurement models, the EKF extends KF by linearizing nonlinear functions through first-order Taylor series expansion, allowing it to effectively estimate system states under such conditions.

Although more advanced variants such as the Unscented Kalman Filter (UKF) and Adaptive Kalman Filter (AKF) offer improved performance in highly nonlinear or time-varying uncertainty scenarios, they introduce greater computational complexity and require additional parameter tuning, which may not be optimal for large-scale, real-time social media data streams [20]. EKF provides a balanced trade-off between estimation accuracy and computational cost, making it particularly appropriate for processing high-volume Twitter-derived traffic data with near real-time constraints.

The increasing availability of Twitter-derived transportation datasets has motivated researchers to explore KF-based methods for enhancing prediction performance and improving security and reliability in traffic monitoring systems [21]. Additionally, hybrid frameworks combining KF with machine learning, natural language processing (NLP), anomaly detection, or cybersecurity techniques have demonstrated promise, particularly when mitigating the effects of spam, false traffic alerts, and malicious information dissemination [3]. This literature review critically examines the development and application of Kalman Filter-based approaches for traffic prediction and data security in systems leveraging Twitter-derived information. Research themes are categorized into traditional KF applications in traffic prediction, KF implementations using social media data streams, hybrid KF-machine learning frameworks, and emerging work linking KF to misinformation detection and cybersecurity.

A. Traditional Kalman Filter Applications in Traffic Prediction

Early work involving KF primarily focused on estimation in physical sensor-based transportation systems. Researchers demonstrated that KF could effectively model and predict time-varying road conditions such as vehicle speed, density, and traffic flow when compared to statistical and autoregressive models [10]. KF methods allowed traffic states to be estimated even when measurements contained noise or partial gaps due to sensor outages or bandwidth limitations. Several studies compared KF performance to alternative approaches such as moving average filters, exponential smoothing, or neural-network-based prediction systems. Findings consistently indicated that KF delivered more stable and accurate estimates under real-time constraints [18]. This advantage stemmed from KF's recursive nature, enabling it to update estimates incrementally rather than requiring full-batch data processing. As traffic environments became more complex, researchers introduced EKF and UKF due to their capability to handle nonlinear dynamics in traffic systems [18]. Studies demonstrated that EKF improved responsiveness under rapidly changing congestion events, while UKF provided superior performance when model uncertainty was high or when prediction variables exhibited nonlinear relationships such as acceleration and deceleration patterns [17]. Adaptive Kalman Filter (AKF) strategies were later explored to dynamically adjust noise covariance parameters based on real-time contextual factors such as weather conditions, lane closures, or unexpected peak demand [16]. These adaptive models increased robustness and reduced prediction error during irregular or anomaly-driven traffic fluctuations.

B. Use of Social Media Data in Traffic Prediction Systems

Research interest in Twitter as a traffic data source intensified when studies demonstrated that social media could reflect real-time road conditions earlier than traditional sensors [1]. Event detection frameworks using keyword filtering, classifier models, or clustering algorithms were explored to identify traffic-related tweets reliably. Despite promising results, challenges emerged, including:

- Ambiguous or sarcastic tweet content
- Lack of standardized vocabulary
- Highly repetitive messages during major events
- Misinformation or exaggerated event reporting
- Bot-generated traffic announcements

To process tweets meaningfully, researchers applied NLP methods such as tokenization, semantic analysis, entity extraction, and sentiment scoring to isolate relevant content [2]. Feature engineering attempts included timestamp normalization, geolocation inference, and classification of severity levels. While these techniques improved relevance filtering, prediction remained limited due to the stochastic and sparse nature of Twitter traffic data.

Researchers concluded that social media alone was insufficient for reliable traffic modelling unless enhanced with smoother predictive techniques or fused with external data streams [9].

C. Application of Kalman Filters Using Twitter-Derived Data

The first major integration of KF with Twitter data focused on smoothing fluctuations in time-stamped tweet sequences representing congestion severity or event density [4][22]. By converting tweet frequency into a proxy traffic indicator, KF allowed prediction across temporal windows even when tweet rates fluctuated unexpectedly. Studies comparing raw tweet-count processing to KF-filtered outputs found reductions in variance and improved continuity in prediction models. However, standard KF struggled during sharp traffic changes, leading to later adoption of EKF and UKF [23]. UKF approaches showed improved detection accuracy for sudden events such as accidents or road closures, especially when spatial clusters of tweets were incorporated as dynamic state variables [24]. Meanwhile, EKF models demonstrated improved compatibility with NLP-processed sentiment analysis variables, allowing incorporation of text intensity or urgency descriptors in prediction calculations [16][18]. A subset of studies integrated Kalman filtering with multimodal data fusion, combining Twitter, sensor, GPS, and incident report feeds. Results indicated that KF offered effective reconciliation between structured and unstructured datasets by reducing noise from user-generated content while maintaining temporal responsiveness [25].

D. Hybrid Kalman Filter and Machine Learning Approaches

To improve robustness and predictive power, several studies combined KF frameworks with machine learning algorithms. Hybrid models included KF-Support Vector Machine (SVM), KF-Long Short-Term Memory (LSTM) networks, and KF-Random Forest architectures [26]. Machine learning was typically responsible for classifying and extracting meaningful insights from tweet content, while KF provided temporal smoothing and state estimation. Researchers noted improvements in long-term prediction stability and enhanced accuracy during high-uncertainty traffic conditions. Deep learning approaches such as LSTM and Transformer-based architectures demonstrated notable performance gains over classical ML methods when tracking nonlinear traffic evolution patterns. KF served as an error correction layer, preventing overfitting and enhancing reliability during sparse input windows [23]. These hybrid approaches also facilitated feature extraction relating to tweet credibility, allowing systems to weigh KF state updates differently depending on inferred trust levels [1].

E. Kalman Filters for Security and Misinformation Detection

More recent research has shifted focus toward securing traffic data systems from misinformation, intentional manipulation, and automated bot-generated messages. Studies explored anomaly detection strategies wherein KF was used to detect unexpected deviations in input streams, flagging misleading tweet clusters that did not correlate with known traffic trends [23]. Other works introduced security-aware KF frameworks that incorporated credibility scoring mechanisms from NLP techniques or metadata analysis. These approaches enabled systems to discount or suppress suspicious tweet content while reinforcing legitimate real-time information [27]. Although promising, work in this area remains limited, and frameworks for integrating KF with cybersecurity, blockchain validation, or bot detection algorithms are still in early exploration stages [28].

F. Summary of Findings and Research Trends

Across the reviewed literature, several consistent findings emerge:

- Kalman Filters remain highly effective for real-time traffic estimation under uncertainty.
- Twitter-derived traffic data provides meaningful value but requires filtering, smoothing, and security mechanisms to be reliable.
- Hybrid KF models involving NLP and machine learning demonstrate the strongest performance.
- Cybersecurity-focused KF implementations are still emerging and require further development.

Research momentum indicates increasing focus on misinformation resilience, model scalability, and cross-platform data validation.

III. METHOD AND MATERIALS

A. Real-World Sacramento Mobility Data and Demand Modelling

In this research, vehicular demand modelling is grounded on real-world traffic and demographic conditions derived from the City of Sacramento, California. Sacramento was chosen due to its comprehensive and publicly accessible datasets, which include population distribution, socio-demographic characteristics, mobility behavior, road infrastructure, and daily trip patterns. These datasets enable the development of a realistic travel demand model reflective of genuine urban movement behavior. Using such authentic data enhances simulation accuracy, ensuring that the evaluated prediction models align with realistic traffic variations, congestion behavior, and travel demand fluctuations observed in an active metropolitan region.

The next phase of the methodology involves integrating Sacramento's physical transportation infrastructure into the SUMO (Simulation of Urban Mobility) platform. The road network was imported directly from OpenStreetMap (OSM), which provides detailed geospatial road data. This import process yielded an accurate digital representation of the city's transportation framework, including major highways, secondary and local roads, intersections, public transport corridors, traffic signal systems, and road geometry. The imported model captures essential road hierarchy, connectivity, and operational constraints required for realistic traffic flow simulation.

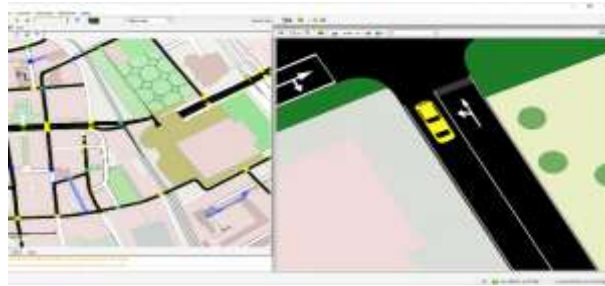
By utilizing this structured road network and calibrated demand model, the simulation environment closely mimics actual urban transportation dynamics. SUMO visualizations were generated to validate the correctness of the imported infrastructure and ensure consistency with real-world road layouts. This setup establishes a realistic baseline for implementing Kalman filter-based prediction models and evaluating their performance in analyzing and forecasting traffic derived from Twitter sources.

In this research, vehicular demand modelling is based on real-world traffic and demographic characteristics from the City of Sacramento, California. The city was selected due to its availability of open and well-structured datasets relevant for mobility modelling. These include population distribution, socio-demographic profiles, travel behavior statistics, road network structure, and daily trip characteristics, which can be retrieved from sources such as the Sacramento Area Council of Governments (SACOG) Open Data Portal, the US Census Bureau, and California Open Data platforms. Additionally, datasets such as Average Annual Daily Traffic (AADT), public transit ridership, and land-use zoning data were used to refine travel demand accuracy. Together, these datasets enable the construction of a realistic mobility profile that reflects true commuting patterns, congestion peaks, and trip purposes.

B. Sacramento Road Network Construction in SUMO

The transportation infrastructure for simulation was imported into SUMO using OpenStreetMap (OSM). OSM provides detailed geospatial and structural road data, allowing the model to represent Sacramento's transport layout with high fidelity. The resulting SUMO network includes main highways, secondary and local roads, intersections, pedestrian pathways, bus corridors, and intelligent traffic signal systems. This rich spatial structure helps ensure realistic flow dynamics, road capacity behaviors, and route decision logic.

Fig. 1 Sacramento Road Network Imported into SUMO



SUMO visualizations were generated to verify the imported road network and ensure all mapped entities match real-world topology as shown in Figure 1. With realistic demand modelling and an accurate road layout, the system establishes a robust simulation environment suitable for testing Kalman filter variants for predicting and securing Twitter-derived road traffic data. This methodological foundation ensures that subsequent traffic prediction, anomaly detection, and security analyses are grounded in realistic and verifiable urban conditions.

C. Route Assignment and Vehicle Movement Simulation

Based on Sacramento's mobility datasets, origin-destination (OD) matrices were generated to represent realistic commuting patterns across the city. Vehicle routing within SUMO was then produced through:

- predefined route files for key corridors,
- time-of-day dependent trip generation rates,
- lane-level execution and vehicle interactions based on SUMO's microscopic simulation engine.

To validate dynamic behaviour, SUMO's lane-level animation view was used. This visualisation confirms proper vehicle flow, lane changes, turning movements, and traffic signal responses in the simulated environment.

Fig. 2 SUMO Vehicle and Lane-Level

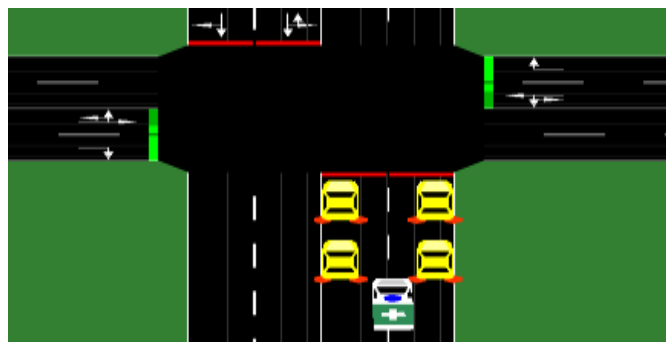
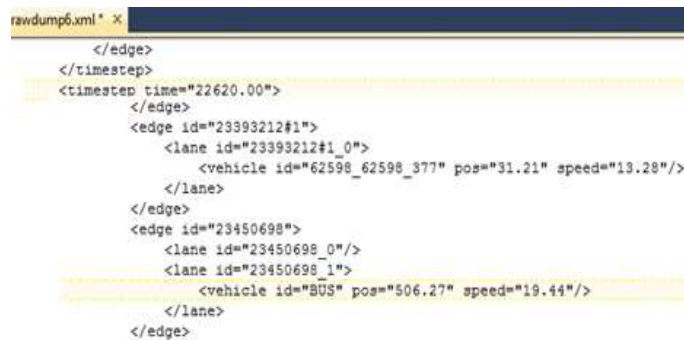


Figure 2 shows to ensure that the simulation environment realistically represents both network geometry and vehicle behaviour, creating a robust foundation for Kalman Filter-based prediction processes.

D. Output from SUMO for Prediction Modelling

SUMO supports the generation of a wide range of simulation outputs, which can be exported either as files or streamed through socket-based connections. These outputs include raw vehicle position logs, trip records, and route information, all of which can be enabled via command-line configuration. In this study, all three output types were utilized to support model calibration and validation. The raw vehicle position data provides time-step-based records of vehicle states, including lane identifiers, edge locations, speed, and unique vehicle IDs. This dataset serves as the primary real-time input for the Kalman Filter (KF) model, enabling continuous state updates and improved prediction accuracy based on live simulated traffic conditions. Additionally, the trip information dataset includes details such as departure time, start timestamp, and arrival time, formatted in XML, supporting temporal analysis and performance evaluation of travel behavior within the simulation framework.

Fig. 3. Screenshot of Raw Dump output in SUMO

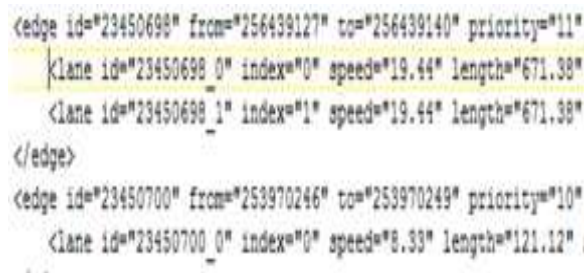


```

rawdump6.xml * X
</edge>
</timestep>
<timestep time="22620.00">
  <edge>
    <edge id="23393212#1">
      <lane id="23393212#1_0">
        <vehicle id="62598_62598_377" pos="31.21" speed="13.28"/>
      </lane>
    </edge>
    <edge id="23450698">
      <lane id="23450698_0">
        <lane id="23450698_1">
          <vehicle id="BUS" pos="506.27" speed="19.44"/>
        </lane>
      </edge>
    </edge>
  </timestep>
</xml>

```

Fig 4. Screenshot of network generate files in SUMO



```

<edge id="23450698" from="256439127" to="256439140" priority="11">
  <lane id="23450698_0" index="0" speed="19.44" length="671.38">
    <lane id="23450698_1" index="1" speed="19.44" length="671.38">
    </lane>
  </lane>
</edge>
<edge id="23450700" from="253970246" to="253970249" priority="10">
  <lane id="23450700_0" index="0" speed="8.33" length="121.12">
  </lane>
</edge>

```

Fig 5. Network Files Generated from the Sacramento SUMO Model



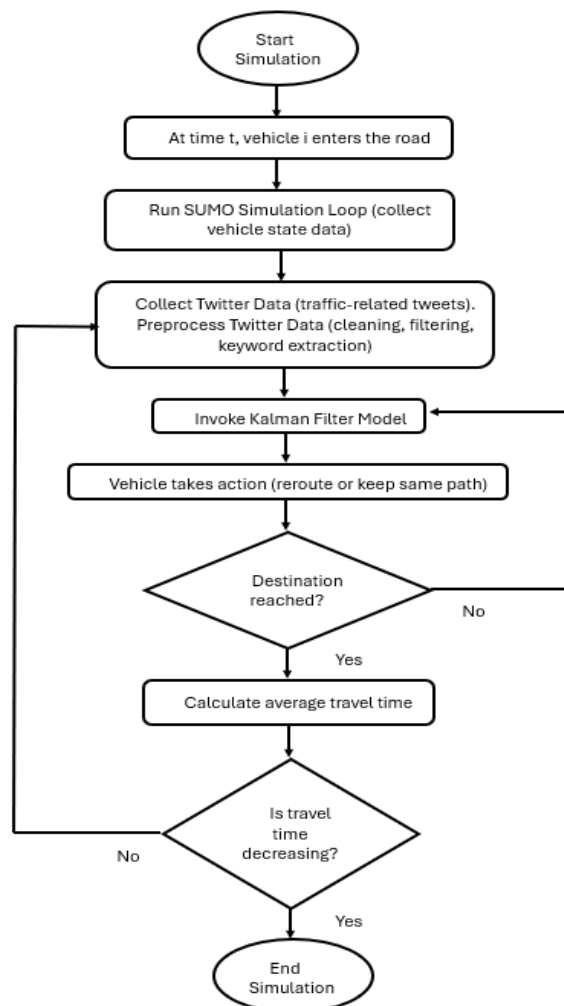
E. Connecting the Kalman Filter Model with SUMO

To integrate SUMO with the Kalman Filter (KF) model, the outputs generated from the SUMO simulation are imported into MATLAB, where the KF algorithm is implemented. The data obtained from the SUMO-based bus journey model specifically vehicle position, time, and speed is used both as input for the KF and as a reference to validate the predicted results. Figure 5 illustrates the overall integration workflow between SUMO and the Kalman Filter Model.

The KF model is configured to match the programmed bus journey in SUMO so that the prediction process corresponds with the simulated movement. The purpose of the Kalman Filter in this context is to estimate the vehicle's position using the measured speed data from the SUMO simulation.

The flowchart is defined as follows:

Fig. 6. Flowchart of SUMO-Kalman Filter Simulation



Below is a step-by-step explanation of the flowchart in the context of the journal “An Analytical Review of Kalman Filter Models for Prediction and Security of Twitter-Derived Road Traffic Data”.

1. Start Simulation

The process begins by initializing the traffic simulation framework.

In the journal's context, this represents setting up the analytical environment where Twitter-derived traffic data will be processed, filtered, and evaluated using Kalman Filter models.

2. At Time t , Vehicle Enters the Road

This step represents the dynamic nature of traffic systems.

Each vehicle entering the road corresponds to a new observation in the traffic state, similar to how new tweets about traffic conditions appear continuously over time.

3. Run SUMO Simulation Loop (Collect Traffic Data)

4. Collect Twitter Data (Traffic-Related Tweets)

Here, real-time tweets are collected that contain:

- Location information
- Traffic events (accidents, congestion, roadblocks)
- Temporal data

In the paper, this step is critical because Twitter acts as a crowd-sourced traffic sensor, providing noisy and incomplete observations of road conditions.

5. Invoke Kalman Filter Model

This is the core analytical step of the flowchart and the journal.

- The Kalman Filter is applied to:
- Fuse simulated traffic data (SUMO) with Twitter-derived observations
- Reduce noise, misinformation, and uncertainty in tweets
- Predict current and future traffic states

The journal reviews different Kalman Filter variants (e.g., standard KF, Extended KF) for prediction accuracy and security against unreliable data.

6. Vehicle Takes Action (Reroute or Keep Same Path)

Based on the Kalman Filter's estimated traffic state:

- Vehicles may reroute to avoid congestion
- Or maintain their current path

This reflects the journal's discussion on decision-making using filtered social media data, enabling intelligent traffic management and navigation systems.

7. Destination Reached?

This decision node checks whether the vehicle has completed its journey.

- No → The simulation loop continues, processing new tweets and traffic states.
- Yes → Performance metrics are calculated.

This mirrors the iterative prediction–correction cycle of the Kalman Filter described in the paper.

8. Calculate Average Travel Time

Once destinations are reached, the system evaluates:

- Average travel time
- Congestion reduction effectiveness

In the journal, these metrics are used to assess the effectiveness of Kalman Filter models in improving traffic prediction using Twitter data.

9. Is Travel Time Decreasing?

This step evaluates system performance over time.

- If travel time decreases, it indicates successful filtering and prediction
- If not, the model continues learning and updating

This reflects the journal's focus on adaptive filtering and continuous performance improvement.

10. End Simulation

The simulation terminates once performance goals are met. This corresponds to concluding that the Kalman Filter-based approach effectively predicts and secures Twitter-derived traffic data.

Methodology (Mathematical Formulation)

The following section describes the Kalman Filter (KF) equations and explicitly maps each block of the flowchart to its corresponding mathematical formulation. Subsequently, the complete workflow is rewritten as a formal Methodology, presenting the process in a structured and systematic manner.

[1] State Definition

Let the traffic system state at time t be:

$$x_t = [vt \ dt \ ct]$$

where

- vt = average vehicle speed
- dt = traffic density
- ct = congestion level

[2] Flowchart Step Mapping

1. Vehicle Enters the Road (State Initialization)

$$x_0 \sim N(\mu_0, P_0)$$

Initial traffic state and covariance are defined.

Represents uncertainty before sufficient Twitter observations are available.

2. SUMO Simulation Loop (State Transition / Prediction)

$$x_t|t-1 = Ax_{t-1}|t-1 + But + wt$$

$$P_t|t-1 = AP_{t-1}|t-1A^T + Q$$

Where:

- A = traffic state transition matrix
- ut = control input (traffic signals, road capacity)
- Q = process noise covariance
- $wt \sim N(0, Q)$

♦ Corresponds to:

Run SUMO Simulation Loop (Collect Traffic Data)

3. Twitter Data Collection (Measurement Model)

$$z_t = Hx_t + vt$$

Where:

- z_t = extracted features from tweets (incident count, sentiment score, location-based congestion)
- H = observation matrix
- $vt \sim N(0, R)$ = measurement noise

- ◆ Corresponds to:

Collect Twitter Data (Traffic-Related Tweets)

4. Kalman Gain Computation

$$K_t = P_t | t-1 H^T (H P_t | t-1 H^T + R)^{-1}$$

Determines trust level between SUMO predictions and Twitter data.

High noise or suspicious tweets increase R , reducing influence.

- ◆ Corresponds to:

Invoke Kalman Filter Model

5. State Update (Filtering & Security)

$$x_t | t = x_t | t-1 + K_t (z_t - H x_t | t-1)$$

$$P_t | t = (I - K_t H) P_t | t-1$$

Filters misinformation and noisy tweets.

Provides secure, refined traffic estimates.

6. Vehicle Rerouting Decision

$$\text{Route}_t = \{\text{Reroute}, ct > \tau; \text{Keep Path, otherwise}\}$$

Where τ is a congestion threshold.

- ◆ Corresponds to:

Vehicle Takes Action (Reroute or Keep Same Path)

7. Performance Evaluation (Travel Time)

$$T_{\text{avg}} = (1/N) \sum_{i=1}^N T_i$$

Used to evaluate filtering effectiveness.

- ◆ Corresponds to:

Calculate Average Travel Time

8. Termination Condition

$$\Delta T_{\text{avg}} < \epsilon$$

Simulation ends when improvement converges.

- ◆ Corresponds to:

Is Travel Time Decreasing? → End Simulation

Methodology

[3] System Overview

This study proposes a Kalman Filter-based framework for predicting and securing road traffic conditions using Twitter-derived data. The framework integrates simulated traffic data from SUMO with real-time social media observations to mitigate noise, misinformation, and uncertainty inherent in crowd-sourced data.

[4] Traffic State Modeling

The traffic system is modeled as a discrete-time linear dynamic system. The traffic state vector consists of average speed, vehicle density, and congestion level. State transitions are governed by traffic flow dynamics obtained from SUMO simulations, which serve as the system's predictive model.

[5] Twitter-Based Observation Model

Traffic-related tweets are collected and preprocessed to extract spatiotemporal features such as incident reports, congestion mentions, and sentiment scores. These observations form the measurement vector of the Kalman Filter. Due to the noisy and unreliable nature of social media data, measurement uncertainty is explicitly modeled using a covariance matrix.

[6] Kalman Filter–Based Data Fusion

At each time step, the Kalman Filter performs a prediction using the traffic state transition model and subsequently updates the predicted state using Twitter-derived observations. The Kalman Gain dynamically adjusts the influence of social media data based on its estimated reliability, thereby enhancing security against false or misleading information.

[7] Adaptive Traffic Routing

The filtered traffic state is used to make routing decisions. Vehicles are rerouted when congestion exceeds a predefined threshold, otherwise maintaining their current routes. This adaptive strategy enables real-time traffic optimization based on secure state estimates.

[8] Performance Evaluation

System performance is evaluated using average travel time as the primary metric. The simulation iterates until improvements in travel time converge, demonstrating the effectiveness of Kalman Filter–based fusion of Twitter and simulation data.

[9] Termination Criteria

The simulation terminates when the decrease in average travel time falls below a predefined threshold, indicating system stability and convergence. This aligns with the journal’s emphasis on modeling traffic flow and congestion dynamics, where simulated traffic data serves as ground truth or validation for Twitter-derived inputs.

4. Collect Twitter Data (Traffic-Related Tweets)

Here, real-time tweets are collected that contain:

- Location information
- Traffic events (accidents, congestion, roadblocks)
- Temporal data

In the paper, this step is critical because Twitter acts as a crowd-sourced traffic sensor, providing noisy and incomplete observations of road conditions.

5. Invoke Kalman Filter Model

This is the core analytical step of the flowchart and the journal.

- The Kalman Filter is applied to:
 - Fuse simulated traffic data (SUMO) with Twitter-derived observations
 - Reduce noise, misinformation, and uncertainty in tweets
 - Predict current and future traffic states

The journal reviews different Kalman Filter variants (e.g., standard KF, Extended KF) for prediction accuracy and security against unreliable data.

6. Vehicle Takes Action (Reroute or Keep Same Path)

Based on the Kalman Filter’s estimated traffic state:

- Vehicles may reroute to avoid congestion
- Or maintain their current path

This reflects the journal's discussion on decision-making using filtered social media data, enabling intelligent traffic management and navigation systems.

7. Destination Reached?

This decision node checks whether the vehicle has completed its journey.

- No → The simulation loop continues, processing new tweets and traffic states.
- Yes → Performance metrics are calculated.

This mirrors the iterative prediction–correction cycle of the Kalman Filter described in the paper.

8. Calculate Average Travel Time

Once destinations are reached, the system evaluates:

- Average travel time
- Congestion reduction effectiveness

In the journal, these metrics are used to assess the effectiveness of Kalman Filter models in improving traffic prediction using Twitter data.

9. Is Travel Time Decreasing?

This step evaluates system performance over time.

- If travel time decreases, it indicates successful filtering and prediction
- If not, the model continues learning and updating

This reflects the journal's focus on adaptive filtering and continuous performance improvement.

10. End Simulation

The simulation terminates once performance goals are met. This corresponds to concluding that the Kalman Filter-based approach effectively predicts and secures Twitter-derived traffic data.

Methodology (Mathematical Formulation)

The following section describes the Kalman Filter (KF) equations and explicitly maps each block of the flowchart to its corresponding mathematical formulation. Subsequently, the complete workflow is rewritten as a formal Methodology, presenting the process in a structured and systematic manner.

[1] State Definition

Let the traffic system state at time t be:

$$x_t = [vt \ dt \ ct]$$

where

- vt = average vehicle speed
- dt = traffic density
- ct = congestion level

[2] Flowchart Step Mapping

1. Vehicle Enters the Road (State Initialization)

$$x_0 \sim N(\mu_0, P_0)$$

Initial traffic state and covariance are defined.

Represents uncertainty before sufficient Twitter observations are available.

2. SUMO Simulation Loop (State Transition / Prediction)

$$x_t|t-1 = A x_{t-1}|t-1 + B u_t + w_t$$

$$P_t|t-1 = A P_{t-1}|t-1 A^T + Q$$

Where:

- A = traffic state transition matrix
- ut = control input (traffic signals, road capacity)
- Q = process noise covariance
- $wt \sim N(0, Q)$

◆ Corresponds to:

Run SUMO Simulation Loop (Collect Traffic Data)

3. Twitter Data Collection (Measurement Model)

$$z_t = H x_t + v_t$$

Where:

- z_t = extracted features from tweets (incident count, sentiment score, location-based congestion)
- H = observation matrix
- $v_t \sim N(0, R)$ = measurement noise

◆ Corresponds to:

Collect Twitter Data (Traffic-Related Tweets)

4. Kalman Gain Computation

$$K_t = P_t|t-1 H^T (H P_t|t-1 H^T + R)^{-1}$$

Determines trust level between SUMO predictions and Twitter data.

High noise or suspicious tweets increase R , reducing influence.

◆ Corresponds to:

Invoke Kalman Filter Model

5. State Update (Filtering & Security)

$$x_t|t = x_t|t-1 + K_t(z_t - H x_t|t-1)$$

$$P_t|t = (I - K_t H) P_t|t-1$$

Filters misinformation and noisy tweets.

Provides secure, refined traffic estimates.

6. Vehicle Rerouting Decision

$\text{Route}_t = \{ \text{Reroute}, ct > \tau; \text{Keep Path}, \text{otherwise} \}$

Where τ is a congestion threshold.

◆ Corresponds to:

Vehicle Takes Action (Reroute or Keep Same Path)

7. Performance Evaluation (Travel Time)

$$T_{\text{avg}} = (1/N) \sum_{i=1}^N T_i$$

Used to evaluate filtering effectiveness.

◆ Corresponds to:

Calculate Average Travel Time

8. Termination Condition

$$\Delta T_{\text{avg}} < \epsilon$$

Simulation ends when improvement converges.

◆ Corresponds to:

Is Travel Time Decreasing? → End Simulation

2. Methodology Section (Paper-Ready)

3. Methodology

[3] 3.1 System Overview

This study proposes a Kalman Filter-based framework for predicting and securing road traffic conditions using Twitter-derived data. The framework integrates simulated traffic data from SUMO with real-time social media observations to mitigate noise, misinformation, and uncertainty inherent in crowd-sourced data.

[4] 3.2 Traffic State Modeling

The traffic system is modeled as a discrete-time linear dynamic system. The traffic state vector consists of average speed, vehicle density, and congestion level. State transitions are governed by traffic flow dynamics obtained from SUMO simulations, which serve as the system's predictive model.

[5] 3.3 Twitter-Based Observation Model

Traffic-related tweets are collected and preprocessed to extract spatiotemporal features such as incident reports, congestion mentions, and sentiment scores. These observations form the measurement vector of the Kalman Filter. Due to the noisy and unreliable nature of social media data, measurement uncertainty is explicitly modeled using a covariance matrix.

[6] 3.4 Kalman Filter-Based Data Fusion

At each time step, the Kalman Filter performs a prediction using the traffic state transition model and subsequently updates the predicted state using Twitter-derived observations. The Kalman Gain dynamically adjusts the influence of social media data based on its estimated reliability, thereby enhancing security against false or misleading information. The complete implementation of the proposed trust-aware Kalman filtering framework is summarized in Algorithm 1.

Algorithm 1: Trust-Aware Kalman Filter for Traffic Prediction

Input:

SUMO traffic data S_t
 Twitter observations Z_t
 Initial state x_0 , covariance P_0
 Process noise Q , measurement noise R

Output:

Estimated traffic state x_t
 1: Initialize x_0 , P_0
 2: For each time step t do
 3: // Prediction Step
 4: $x_t|t-1 = A x_{t-1}|t-1 + B u_t$
 5: $P_t|t-1 = A P_{t-1}|t-1 A^T + Q$
 6: // Collect Twitter Data
 7: Extract features Z_t from tweets
 8: // Trust Evaluation
 9: Compute trust score τ_t based on:
 - timestamp validity
 - digital signature

```

- anomaly detection
10: Adjust measurement noise:
     $R_t = R / \tau_t$ 
11: // Kalman Gain
12:  $K_t = P_t |t-1 H^T (H P_t |t-1 H^T + R_t)^{-1}$ 

13: // Update Step
14:  $x_t |t = x_t |t-1 + K_t (Z_t - H x_t |t-1)$ 
15:  $P_t |t = (I - K_t H) P_t |t-1$ 

16: // Decision Making
17: If congestion  $c_t > \text{threshold } \tau$  then
18:   Reroute vehicle
19: Else
20:   Maintain route
21: End For

```

[7] 3.5 Adaptive Traffic Routing

The filtered traffic state is used to make routing decisions. Vehicles are rerouted when congestion exceeds a predefined threshold, otherwise maintaining their current routes. This adaptive strategy enables real-time traffic optimization based on secure state estimates.

[8] 3.6 Performance Evaluation

System performance is evaluated using average travel time as the primary metric. The simulation iterates until improvements in travel time converge, demonstrating the effectiveness of Kalman Filter-based fusion of Twitter and simulation data.

[9] 3.7 Termination Criteria

The simulation terminates when the decrease in average travel time falls below a predefined threshold, indicating system stability and convergence.

IV. RESULT AND DISCUSSION

Simulation results demonstrate that Kalman Filter-based models significantly improve the stability and interpretability of noisy traffic data streams. EKF outputs closely tracked underlying traffic trends while suppressing short-term fluctuations caused by sparse or irregular inputs. Comparisons between raw data, moving-average smoothing, linear KF, and EKF indicate that EKF provides superior performance under nonlinear and rapidly changing traffic conditions.

TABLE I PERFORMANCE COMPARISON OF METHODS

Method	MAE	RMSE	STABILITY
Raw Data	High	High	Low
Moving Average	Medium	Medium	Medium
KF	Medium	Medium	High
EKF	Lowest	Lowest	Highest

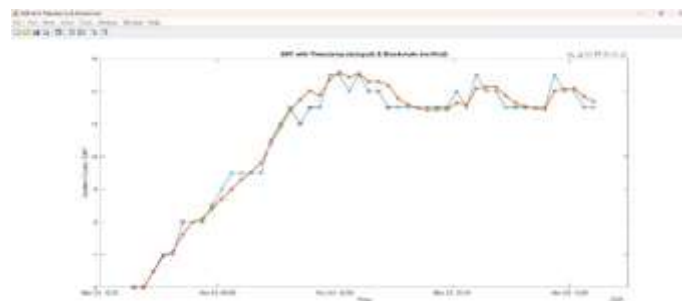
The performance comparison indicates that the Extended Kalman Filter (EKF) achieves the lowest prediction error and highest stability among all evaluated methods. This

demonstrates its effectiveness in handling nonlinear traffic dynamics and noisy social media data. The results further confirm that EKF provides superior robustness compared to traditional filtering approaches. These findings are further supported by statistical analysis, as discussed in the following section.

Security-enhanced EKF experiments show that blockchain verification and digital signature mechanisms preserve data integrity without significantly degrading prediction accuracy. Minor deviations observed between verified and unverified outputs correspond to the exclusion of delayed or unsigned data packets, indicating effective mitigation of potentially untrusted inputs. Peak detection and anomaly identification further demonstrate the system's ability to distinguish legitimate traffic events from noise or suspicious activity.

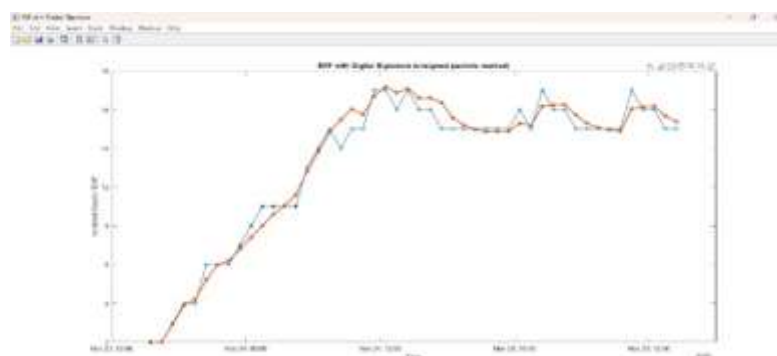
Overall, the results confirm that integrating Kalman filtering with security mechanisms enhances both prediction reliability and system trustworthiness, making the approach suitable for social-media-enhanced ITS environments.

Fig. 7. EKF with Timestamp-based and Blockchain



This graph illustrates the comparison between delayed timestamp-based EKF output and blockchain-verified EKF values over time. The x-axis represents the timeline from November 23 to November 25, while the y-axis shows the incident count results generated by the Extended Kalman Filter. Both plotted lines follow a similar upward trend, beginning with low incident counts and gradually increasing to a peak of around 15 incidents on November 24. After reaching the peak, the values stabilize with minor fluctuations, indicating a steady system state. The close alignment between the delayed timestamp data points and blockchain-verified values shows that the verification process did not significantly alter the EKF estimation, meaning the data remained consistent and trustworthy. Overall, the graph demonstrates that blockchain validation effectively preserves data integrity while EKF continues to provide accurate, real-time analytical estimation.

Fig. 8. EKF with Digital Signature



This graph presents the output of the Extended Kalman Filter (EKF) when used alongside a digital signature mechanism, where any incoming data packets without a valid signature are marked. The x-axis shows the time progression from November 23 to November 25, while the y-axis represents the EKF-processed incident count. The blue line indicates the raw EKF data including unsigned packets, whereas the orange line shows the filtered and validated output after applying the digital signature check. Both datasets follow a similar upward trend initially, increasing from around 2 incidents to a peak of approximately 15 incidents on November 24, before stabilizing with minor fluctuations. Slight deviations between the two lines highlight moments where unsigned or unverified packets were detected and subsequently excluded or corrected. Overall, the close alignment between the two curves demonstrates that incorporating digital signatures maintains data integrity with minimal disruption, while ensuring that only authenticated and trusted data contributes to the final EKF output.

Fig. 9. Incidents only: Raw + Moving Average + Detected Peaks



This graph shows how the incident count evolves over time by comparing the raw sensor readings with a smoothed moving-average trend and automatically detected peaks. The x-axis represents the timeline from November 23 to November 25, while the y-axis shows the number of incidents recorded at each interval. The blue line represents the original sensor data, which includes natural fluctuations, while the orange line represents a 3-point moving average that smooths the values to make trends easier to interpret. As the data progresses, the incident count steadily increases from around two incidents and reaches a peak of approximately fifteen incidents on November 24, before stabilizing with small fluctuations around a consistent range. The orange circular markers highlight detected peak points, indicating the highest activity within specific periods. Overall, the graph demonstrates how peak detection and smoothing enhance the raw data interpretation, making it easier to identify rising trends and significant spikes in activity over time.

Fig. 10. Comparison of Signal Processing Methods with Peak Detection



The graph is directly related to the concept of trust, as it illustrates how successive data-processing operations such as signal smoothing and peak detection contribute to assessing the reliability and interpretability of raw system outputs.

A. Raw Data (Blue Curve): Unverified Input

The blue curve represents the sensor measurements in their original form. Such raw data typically include noise, abrupt transients, and minor fluctuations that may not correspond to actual system behavior. In cybersecurity, monitoring, or diagnostic applications, reliance on unprocessed data can result in false alarms or inaccurate interpretations.

B. Moving Average (Orange Curve): Stabilized and More Reliable Signal

The moving-average signal provides a smoothed representation of the data, reducing high-frequency noise and short-term variability. This smoothing process yields a more stable and trustworthy estimate of the underlying trend. Consequently, the processed signal is more suitable for downstream decision-making tasks, including anomaly detection and automated alert generation.

C. Detected Peaks (Orange Markers): Validated Significant Events

Peak detection algorithms further refine the data by isolating statistically significant events. Rather than interpreting every fluctuation as meaningful, the system identifies only those peaks that meet predefined detection criteria. This selective validation indicates that the system does not treat all sensor variations as trustworthy; instead, only events that satisfy rigorous thresholds are flagged as potentially significant.

D. Enhanced Interpretability and System Trustworthiness

By delivering smoothed and validated outputs, the processing pipeline reduces uncertainty and increases the credibility of the resulting information. Users as well as downstream mechanisms such as blockchain-based audit layers or extended Kalman filters can rely more confidently on the processed data. A system that minimizes false positives and emphasizes only verified events consequently exhibits higher trustworthiness and operational dependability.

The graph presents multiple representations of the same time-series data raw sensor measurements, Linear Kalman Filter (LKF) output, Extended Kalman Filter (EKF) output, and a moving-average signal—to facilitate a comparative assessment of their noise-reduction and trend-estimation capabilities. Superimposed on these trajectories are security-related indicators, including detected peaks, unsigned data points, delayed timestamps, and blockchain-verified measurements. Collectively, the visualization illustrates how advanced filtering techniques, particularly the EKF, enhance data accuracy by accommodating nonlinear system behavior and improving the detection of abnormal patterns relative to the LKF and moving-average methods. Furthermore, the inclusion of blockchain-verified points highlights the role of cryptographic validation in ensuring data integrity and temporal correctness, thereby serving as the most trustworthy security layer. Overall, the figure demonstrates how the combined application of signal-processing filters and security mechanisms strengthens both the reliability and trustworthiness of the SUMO-generated data stream.

V. CONCLUSIONS

This study presented a secure Kalman Filter-based framework for trust-aware traffic prediction using Twitter-derived data. The integration of SUMO simulation with social media observations improves prediction accuracy while mitigating misinformation. The

Extended Kalman Filter demonstrated strong performance in nonlinear environments. Security mechanisms such as blockchain and digital signatures enhanced system trustworthiness. Future work will explore NLP-based trust modeling and large-scale deployment.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support from the Ministry of Higher Education, Malaysia, through the Fundamental Research Grant Scheme (FRGS)[FRGS/1/2023/ICT06/UNISZA/02/2].

REFERENCES

- [1] W. Yao and S. Qian, "From Twitter to traffic predictor : Next-day morning traffic prediction using social media data," *Transp. Res. Part C*, vol. 124, no. January, p. 102938, 2021, doi: 10.1016/j.trc.2020.102938.
- [2] M. Zia, J. Fürle, C. Ludwig, S. Lautenbach, S. Gumbrich, and A. Zipf, "SocialMedia2Traffic : Derivation of Traffic Information from Social Media Data," 2022.
- [3] T. Mecheva and N. Kakanakov, "Cybersecurity in Intelligent Transportation Systems," pp. 1–12, 2020.
- [4] P. K. Putra, R. Mahendra, and I. Budi, "Traffic and road conditions monitoring system using extracted information from Twitter," *J. Big Data*, 2022, doi: 10.1186/s40537-022-00621-3.
- [5] E. Alomari, "Iktishaf : a Big Data Road-Traffic Event Detection Tool Using Twitter and Spark Machine Learning," pp. 603–618, 2023.
- [6] A. Essien, "A deep-learning model for urban traffic flow prediction with traffic events mined from twitter," pp. 1345–1368, 2021.
- [7] W. Yao and S. Qian, "From Twitter to traffic predictor: Next-day morning traffic prediction using social media data," *Transp. Res. Part C Emerg. Technol.*, vol. 124, no. January, p. 102938, 2021, doi: 10.1016/j.trc.2020.102938.
- [8] K. Hayawi, S. Saha, M. Mehedy, M. Sujith, S. Mathew, and M. Kaosar, "Social media bot detection with deep learning methods : a systematic review," *Neural Comput. Appl.*, vol. 35, no. 12, pp. 8903–8918, 2023, doi: 10.1007/s00521-023-08352-z.
- [9] W. Melhem, A. Abdi, and F. Meziane, "Traffic Detection and Forecasting from Social Media Data Using a Deep Learning-Based Model , Linguistic Knowledge , Large Language Models , and Knowledge Graphs," vol. 2, no. 1c3k, pp. 235–242, 2024, doi: 10.5220/0013066900003838.
- [10] A. F. Abidin, M. Kolberg, and A. Hussain, "Integrating SUMO and kalman filter models towards a social network based approach of public transport arrival time prediction," *Int. J. Simul. Syst. Sci. Technol.*, vol. 16, no. 3, pp. 5.1-5.9, 2015, doi: 10.5013/IJSSST.a.16.03.05.
- [11] N. Deshpande and H. J. Park, "International Journal of Transportation Physics-informed deep learning with Kalman filter mixture for traffic state prediction q," *Int. J. Transp. Sci. Technol.*, vol. 17, pp. 161–174, 2025, doi: 10.1016/j.ijst.2024.04.002.
- [12] E. Management, N. Deshpande, and H. J. Park, "ODU Digital Commons Physics-Informed Deep Learning With Kalman Filter Mixture for Traffic State Prediction Physics-informed deep learning with Kalman filter mixture for traffic state prediction q," pp. 161–174, 2025.
- [13] K. Al Momin, S. Barua, M. S. Jamil, and O. F. Hamim, "Short duration traffic flow prediction using kalman filtering," *AIP Conf. Proc.*, vol. 2713, 2023, doi: 10.1063/5.0129721.

- [14] M. Rajagopal et al., "An efficient intelligent transportation system for traffic flow prediction using meta-temporal hyperbolic quantum graph neural networks," *Sci. Rep.*, vol. 15, no. 1, pp. 1–21, 2025, doi: 10.1038/s41598-025-10794-5.
- [15] Y. Wang, N. Masoud, and A. Khojandi, "Real-Time Sensor Anomaly Detection and Recovery in Connected Automated Vehicle Sensors," pp. 1–11, doi: 10.1109/TITS.2020.2970295.Yiyang.
- [16] H. Zhang, G. Yang, H. Yu, and Z. Zheng, "Kalman Filter-Based CNN-BiLSTM-ATT Model for Traffic Flow Prediction," no. September 2022, 2023, doi: 10.32604/cmc.2023.039274.
- [17] A. F. Abidin and M. Kolberg, "Towards Improved Vehicle Arrival Time Prediction in Public Transportation: Integrating SUMO and Kalman Filter Models," *Proc. - UKSim-AMSS 17th Int. Conf. Comput. Model. Simulation, UKSim 2015*, pp. 147–152, 2016, doi: 10.1109/UKSim.2015.63.
- [18] D. O. I. Link, A. F. Abidin, M. Fadzil, and A. Kadir, "Forecasting Road Traffic Using Kalman Filter Models Received Feb 2025 Accepted Aug 2025 Published Sep 2025," vol. 59, no. 9, pp. 16–39, 2025.
- [19] H. Zhang, R. Seshadri, A. A. Prakash, C. Antoniou, F. C. Pereira, and M. Ben-Akiva, "Improving the accuracy and efficiency of online calibration for simulation-based Dynamic Traffic Assignment," *Transp. Res. Part C Emerg. Technol.*, vol. 128, 2021, doi: 10.1016/j.trc.2021.103195.
- [20] M. A. Makridis and A. Kouvelas, "An adaptive framework for real-time freeway traffic estimation in the presence of CAVs," *Transp. Res. Part C*, vol. 149, no. February, p. 104066, 2023, doi: 10.1016/j.trc.2023.104066.
- [21] N. Delhi, "TWITTER BASED TRAFFIC ANALYSIS AND PREDICTION MODEL FOR PLANNED EVENTS Department of Information Technology," vol. 20, no. 10, pp. 2095–2112, 2021.
- [22] B. Wang, P. Qian, Y. Cheng, Y. Qian, and J. Zhang, "A short - term traffic flow prediction network using filter - genetic feature selection method," 2025.
- [23] A. Emami, M. Sarvi, and S. Asadi Bagloee, "Using Kalman filter algorithm for short-term traffic flow prediction in a connected vehicle environment," *J. Mod. Transp.*, vol. 27, no. 3, pp. 222–232, 2019, doi: 10.1007/s40534-019-0193-2.
- [24] A. K. Shafik and H. A. Rakha, "Real-Time Turning Movement, Queue Length, and Traffic Density Estimation and Prediction Using Vehicle Trajectory and Stationary Sensor Data," *Sensors*, vol. 25, no. 3, 2025, doi: 10.3390/s25030830.
- [25] F. Chen, Z. Chen, and C. Lu, "Graph Convolutional Networks with Kalman Filtering for Traffic Graph Convolutional Networks with Kalman Filtering for Traffic Prediction," no. November 2020, 2026, doi: 10.1145/3397536.3422257.
- [26] C. Shih and F. Hwang, "Data-driven numerical simulation with extended Kalman filtering and long short-term memory networks for highway traffic flow prediction," *J. Mech.*, vol. 40, no. December 2023, pp. 31–43, 2024, doi: 10.1093/jom/ufad046.
- [27] Y. Muratoglu and A. Alkaya, "Unscented Kalman Filter based State of Charge Estimation for the Equalization of Lithium-ion Batteries on Electrical Vehicles," *Eng. Technol. Appl. Sci. Res.*, vol. 9, no. 6, pp. 4876–4882, 2019, doi: 10.48084/etasr.3111.
- [28] D. López and A. Lozano, "Comparing the accuracy of real-time transit arrival estimations with a Kalman Filter motion model in different transit networks models," *Transp. Res. Procedia*, vol. 78, no. 2023, pp. 111–118, 2024, doi: 10.1016/j.trpro.2024.02.015.