LSTM-based IDS for VANETs: A Time Series Classification Approach to False Message Detection

Yantao Yu

Contact Email: yantao.yu@unb.ca

Canadian Institute for Cybersecurity (CIC), Faculty of Computer Science, University of New Brunswick (UNB)

ABSTRACT

In vehicular ad hoc networks (VANETs), vehicles broadcast emergency messages (EMs) and beacon messages (BMs), which enable drivers to perceive traffic conditions beyond their visual range thus improve driving safety. However, internal attackers can launch a false message attack for selfish purposes by reporting a non-existent traffic incident in EMs. Moreover, some collusion attackers may spread bogus BMs cooperatively to make the bogus traffic incident more deceptive. To improve the accuracy of false EM detection, we propose a novel intrusion detection system (IDS) based on time series classification and deep learning. Considering that traffic parameters are highly correlated with time, we collect time series of traffic parameters closely related to traffic incidents from messages of vehicles near reported traffic incidents as time series feature vectors. To recognize the pattern of traffic parameters changing over time more accurately, a traffic incident classifier based on long short-term memory (LSTM) is designed and trained using time series feature vectors from both normal and collusion attack scenarios. Based on the classification result, the authenticity of the EM can be determined.

System Model

(1) VANETs Model

- **Own vehicle:** The considered vehicle that receives an EM and is about to run the IDS to detect false messages.
- **Warning vehicle:** The witness of the traffic incident and the EM sender.
- **Neighboring vehicles:** One-hop neighbors of the own vehicle.

Each vehicle broadcasts a BM every time t_{b}. The own vehicle (V_0) stores BMs of neighboring vehicles (V_1 to V_6) for the latest observation time t_{ob} in the Current Neighbor List (CNL). Suppose that at time t_0, V_1 senses a traffic incident, and it will broadcast an EM to vehicles behind. When V_0 receives the EM, its CNL stores BMs of neighboring vehicles within$[t_0, t_0\text{-}t_{\text{ob}}]$. To verify the authenticity of the EM, V_0 collects evidence related to vehicles in the observation area (V_1, V_2, and V_3) from CNL and runs the IDS. The observation area is near the warning vehicle and is included in the communication range of the own vehicle.

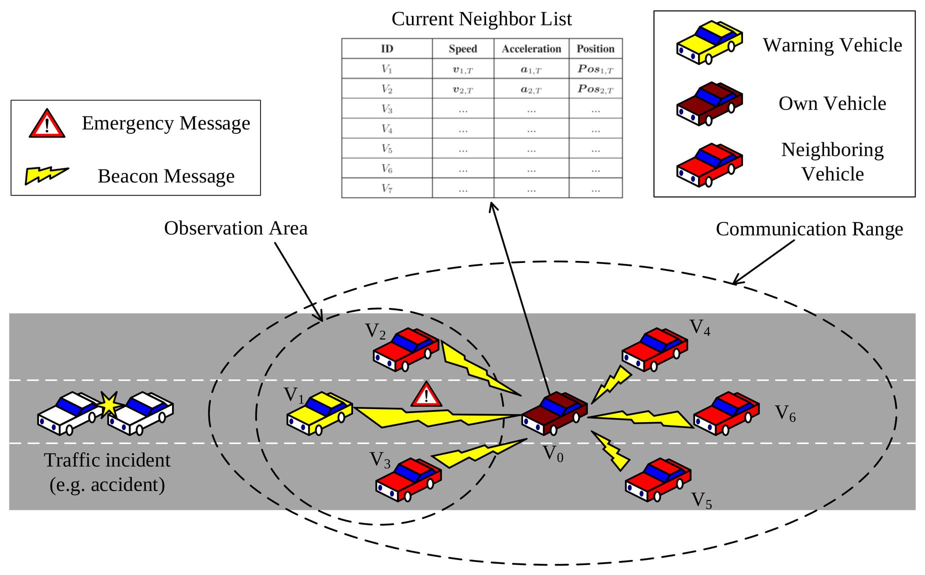

Fig. 1. VANETs model on a highway

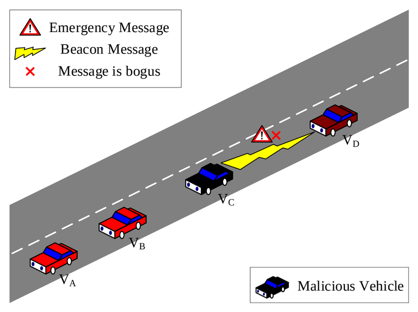

Fig. 2. False EM attack

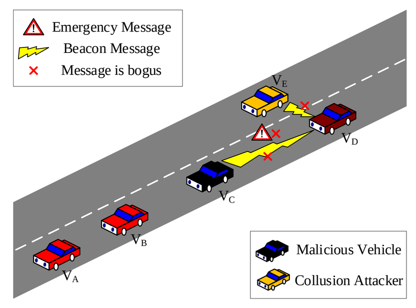

Fig. 3. Collusion attack

(2) Message Format

- **Beacon Message:** $BeaconMsg(ID, v, a, Pos)$,

where ID, v, a, and Pos are the vehicle's identity, speed, acceleration, and position, respectively.

- **Emergency Message:** $EmergencyMsg(ID, E_{\text{type}}, E_{\text{pos}})$,

where E_{type} is the traffic incident type, such as accidents, poor road conditions, congestion, etc., and E_{pos} is the location of the incident.

(3) Attack Model

- **False Emergency Message Attack:** Declaring a false traffic incident in the EM. Denoted as $EmergencyMsg'(ID, E'_{\text{type}}, E'_{\text{pos}})$.
- **Collusion Attack:** some collusion attackers forge their BMs to simulate the real movement state after encountering a traffic incident. Denoted as $BeaconMsg'(ID, v', a', Pos')$.

Overview of the proposed IDS

Based on the fact that different traffic incidents will cause traffic parameters to fluctuate in different patterns, the core idea of the scheme proposed is as follows.

(1) Collect the traffic parameters within a period of observation time from messages of vehicles near reported traffic incidents to form a multi-dimensional time series (MTS).

(2) Classify the MTS to determine the current traffic incident.

(3) EM's authenticity is judged according to whether the classification result is consistent with the reported traffic incident.

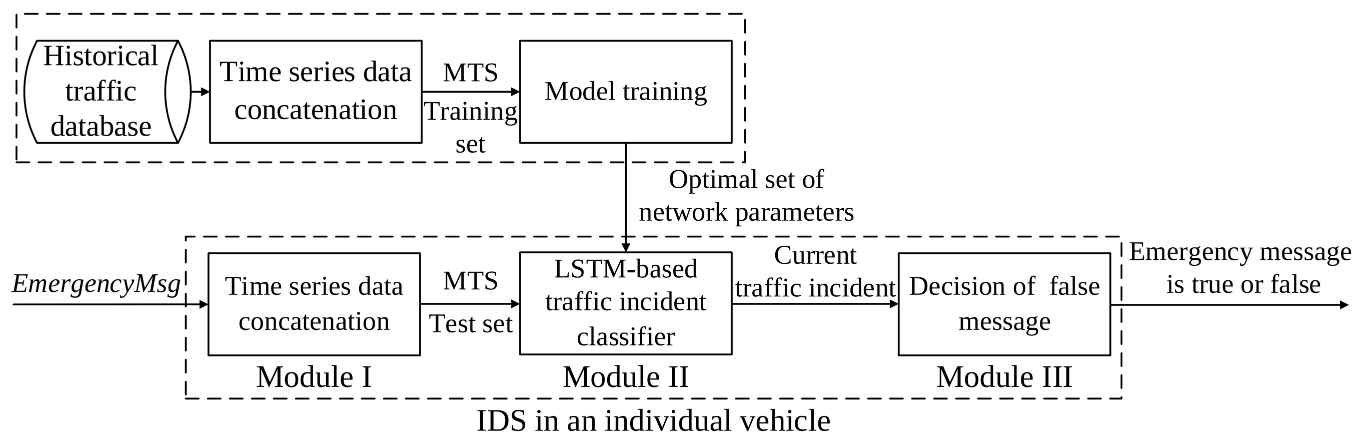

Fig. 4. Workflow of the proposed IDS

A. Time Series Data Concatenation

(1) Time series data storage

Each vehicle maintains the CNL by collecting neighboring vehicles' BMs in the latest period of time t_{ob}, that is, there are data at $n = \lfloor t_{\text{ob}}/t_{\text{b}} \rfloor$ points in time.

Tabel 1. CNL stored by a vehicle at t

ID	Speed	Acceleration	Position
V_1	$v_{1,T}$	$a_{1,T}$	$Pos_{1,T}$
V_2	$v_{2,T}$	$a_{2,T}$	$Pos_{2,T}$
…	…	…	…

$v_{i,T}$ is the time series of the speed of V_i on $T \in \{t - n + 1, \dots, t-1, t\}$, i.e. $v_{i,T} = [v_{i,t-n+1}, \dots, v_{i,t-1}, v_{i,t}]$. The sampling time interval of the data is t_{b}.

The own vehicle V_o receives the EM from V_w (i.e. the warning vehicle) at time t, and computes the time series feature vector (TSFV) from its CNL.

$$\mathbf{X}_{o,T} = [\boldsymbol{v}_{w,T}, \boldsymbol{a}_{w,T}, \boldsymbol{d}_{w,T}, \overline{\boldsymbol{v}}_{w,T}, \overline{\boldsymbol{a}}_{w,T}, \boldsymbol{N}_{w,T}]$$

Tabel 2. The meaning of each feature of TSFV

Feature symbol	Meaning
$\boldsymbol{v}_{w,T}$	The speed of the waring vehicle
$\boldsymbol{a}_{w,T}$	The acceleration of the waring vehicle
$\boldsymbol{d}_{w,T}$	The distance between the warning vehicle and where the incident occurred
$\overline{\boldsymbol{v}}_{w,T}$	The average speed of vehicles in the observation area
$\overline{\boldsymbol{a}}_{w,T}$	The average acceleration of vehicles in the observation area
$\boldsymbol{N}_{w,T}$	The total number of vehicles in the observation area

Note: All features are time series on observation time T

(2) Time series feature vector computation

- Let $S_{\text{nei},t}$ be V_o's neighboring vehicles set and $S_{\text{obs},t}$ be the set of vehicles in the observation area at t, i.e. $S_{\text{obs},t} = \{V_i | V_i \in S_{\text{nei},t}, Dist(Pos_{i,t}, Pos_{w,t}) \le r_{\text{obs}}\}$
- Since $\mathbf{X}_{o,T}$ includes the average traffic parameters $(\overline{\boldsymbol{v}}_{w,T}, \overline{\boldsymbol{a}}_{w,T})$ and they might be dominated by a small number of vehicles who spread extreme data in their BMs, Grubbs's test is used to detect set of outlier vehicles with large-scale modifications of speed data (denoted as S_{out}).

For the computation of $\mathbf{X}_{o,T}$, we can easily get $\boldsymbol{v}_{w,T}$ and $\boldsymbol{a}_{w,T}$ from the CNL of V_o, and $\boldsymbol{d}_{w,T}, \overline{\boldsymbol{v}}_{w,T}, \overline{\boldsymbol{a}}_{w,T}, \boldsymbol{N}_{w,T}$ can be respectively computed at each time point as follows:

- $\boldsymbol{v}_{w,T}$ and $\boldsymbol{a}_{w,T}$ can be easily obtained from CNL
- $d_{w,t} = Dist(Pos_{w,t}, E_{\text{pos}})$,
- $\bar{v}_{w,t} = \sum_{V_i \in S'_{\text{obs},t}} \frac{v_{i,T}}{|S'_{\text{obs},t}|}$,
- $\bar{a}_{w,t} = \sum_{V_i \in S'_{\text{obs},t}} \frac{v_{i,T}}{|S'_{\text{obs},t}|}$,
- $N_{w,T} = |S'_{\text{obs},t}|$,

where $S'_{\text{obs},t} = S_{\text{obs},t} - S_{\text{out}}$

B. LSTM-based Traffic Incident Classifier

(1) Dataset Format

The dataset $D = \{(\mathbf{X}_{1,T}, \hat{\boldsymbol{y}}_1), (\mathbf{X}_{2,T}, \hat{\boldsymbol{y}}_2), \dots, (\mathbf{X}_{N,T}, \hat{\boldsymbol{y}}_N)\}$ is a collection of pairs $(\mathbf{X}_{i,T}, \hat{\boldsymbol{y}}_i)$, where $\mathbf{X}_{i,T}$ is the TSFV, $\hat{\boldsymbol{y}}_i$ is the one-hot label vector of each traffic incident, and N is the total samples number. The size of the $\mathbf{X}_{i,T}$ is $[6 \times n]$. We consider 4 types of common traffic conditions (i.e. normal, accident, poor road condition, and congestion). So $\hat{\boldsymbol{y}}_i$ is a vector with 4 dimensions.

(2) LSTM Network Structure

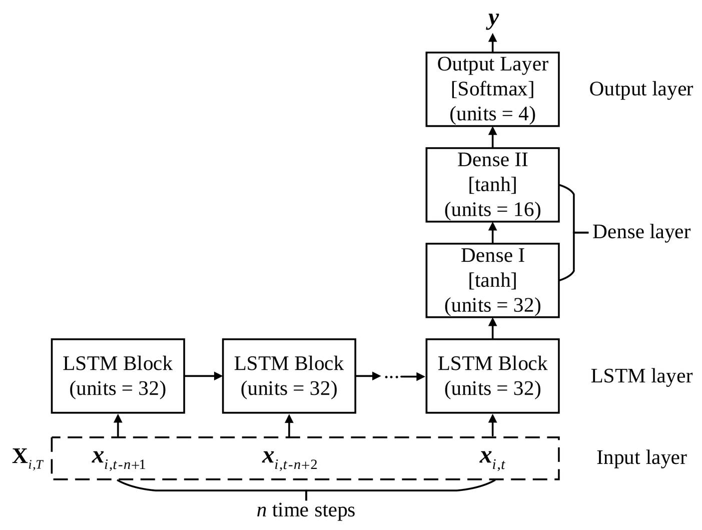

Fig. 5. LSTM Network Structure

(3) Model Training

The loss function can be calculated by:

$$L = \frac{1}{N} \sum_{i=1}^{N} \sum_{k=0}^{3} \left(\hat{\boldsymbol{y}}_i^{(k)} \ln \boldsymbol{y}_i^{(k)} \right) + \frac{\lambda}{2N} \sum_{w} w^2 ,$$

where $\hat{\boldsymbol{y}}_i^{(k)}$ and $\boldsymbol{y}_i^{(k)}$ are the k-th dimension of the i-th sample of the output vector and label vector respectively, w is the connection weights of the network, and λ is the regularization parameter.

C. Decision of False Message

Suppose the result of traffic incident classification is E, and the reported traffic incident is $\hat{E}$. The match factor is calculated by: $x_{\text{match}} = E \oplus \hat{E}$.

- If x_{match} is 0, EM is judged as a real message;
- Otherwise, EM is rejected, and V_w is regarded as a malicious node.

* This work has been accepted by IEEE Transactions on Intelligent Transportation System in 2022.