

10-4-2022

Location verification for future wireless vehicular networks: Research directions and challenges

Shihao Yan
Edith Cowan University, s.yan@ecu.edu.au

Ullah Ihsan

Robert Malaney

Linlin Sun

Stefano Tomasin

Follow this and additional works at: <https://ro.ecu.edu.au/ecuworks2022-2026>



Part of the [Computer Sciences Commons](#)

[10.1109/MNET.103.2100338](https://doi.org/10.1109/MNET.103.2100338)

This is an Authors Accepted Manuscript version of an article published by IEEE in *IEEE Network*, at <https://doi.org/10.1109/MNET.103.2100338>

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Yan, S., Ihsan, U., Malaney, R., Sun, L., & Tomasin, S. (2022). Location verification for future wireless vehicular networks: Research directions and challenges. *IEEE Network*, 36(6), 60-66. <https://doi.org/10.1109/MNET.103.2100338>

This Journal Article is posted at Research Online.
<https://ro.ecu.edu.au/ecuworks2022-2026/1970>

Location Verification for Future Wireless Vehicular Networks: Research Directions and Challenges

Shihao Yan, *Member, IEEE*, Ullah Ihsan, *Student Member, IEEE*, Robert Malaney, *Senior Member, IEEE*,
Linlin Sun, *Member, IEEE*, and Stefano Tomasin, *Senior Member, IEEE*

Abstract—Vehicle location information obtained through the global navigation satellite system (GNSS) will play a pivotal role in emerging vehicular networks. This vital information is, however, susceptible to a host of unwanted manipulations, especially if a malicious entity is involved. The most obvious example of such manipulations is the forwarding by a malicious vehicle of false GNSS locations to other members of the network. Such events can lead to poor operational outcomes for the vehicular network, and in extreme cases even lead to catastrophic safety violations. Here, we highlight research efforts pursued in the past few years which have attempted to address this weakness in vehicular networks. We also discuss the importance of location verification in the wake of emerging wireless technologies such as those being proposed for beyond fifth generation (B5G) wireless vehicular networks. In particular, we detail an opportunity to conduct location reporting and verification simultaneously with the aid of mmWave technology and discuss how emerging machine learning (ML) techniques will provide for location verification solutions whose reliability levels will be commensurate with that required by the vehicular network paradigm. We close by discussing the potential enhancements for location verification within a future combined B5G-ML architecture.

I. INTRODUCTION

A wireless vehicular network (WVN) is a particular type of intelligent transportation system that utilizes vehicle-to-everything (e.g., inter-vehicle and infrastructure-vehicle) communications to carry out different network operations. In the context of communicating cars, with the aid of a WVN we can optimize traffic routing, minimize traffic congestions, improve road tolling infrastructure, assist smart city traffic planning, aid in road infrastructure expansions, achieve seamless in-vehicle entertainment services, increase electric vehicle charging performance [1], and most importantly enhance road safety. As per a survey conducted by the world health organization, there are annually over one million casualties worldwide due to car accidents. Overcoming such a high number of fatalities, even partially, is a challenge that motivates an increasing amount of research on WVN. In order to enable the full functions of a WVN, we require the following three desirable features; ultra-high data rates, ultra-reliability, and ultra-low latency. These desirable features cannot be fully met by existing general wireless communication systems such as fourth-generation

(4G) cellular systems or the dedicated short-range communication (DSRC) system. However, the forthcoming beyond fifth generation (B5G) technology is specifically designed with these desirable features in mind.

Location information plays an important role in WVN and even serves a foundation role in several key functional areas [2]–[4]. For example, many network operations of WVN, such as location-based routing, depend critically on the location information of each vehicle in the network. Also, location information serves as the enabler of many services provided by WVN, e.g., collision avoidance and location-based advertising. The WVN paradigm normally assumes that positioning systems are client-based, i.e., the vehicle itself obtains its location information, normally via an on-board GNSS device. However, these devices are vulnerable to attacks (or faults) and thus may provide fake (or unreliable) location information to the WVN. That is, the reliability and correctness of a vehicle's reported location information cannot be fully guaranteed in this context. An emerging technology that addresses this important issue is a location verification system (LVS). Such systems have attracted an increasing amount of research interest in recent years, e.g. [5].

Different from positioning systems, an LVS aims at confirming whether a user (e.g., a vehicle) is physically at its reported/claimed location. Mathematically, location verification is a binary detection problem, while localization is an estimation problem. The importance of an LVS can be evidenced by the negative effects of fake location information on WVN, e.g., dramatically reduced packet delivery in position-based routing protocols or even life-threatening road accidents. In principle, an LVS can remove these negative impacts on WVN operations. We also note that verified location information can improve the performance of mmWave-based communication (a key enabler of the B5G wireless networks), since this location information can aid in the determination of accurate channel state information (CSI). In order to fully support the services of emerging WVN, location verification should be embedded into emerging B5G technology - a belief that forms the thrust of this article.

In the following we detail the integration of LVSs into WVN via B5G technology, and outline the opportunities this delivers. To this end, we first review location verification frameworks and techniques, and present the state-of-the-art LVS algorithms. We then provide insight on the optimal use of the B5G technology to address location verification. We also identify an opportunity of using mmWave to incorporate location verification into a communication system with a

S. Yan is with Edith Cowan University, Perth, WA 6027, Australia (e-mail: s.yan@ecu.edu.au). U. Ihsan and R. Malaney are with the University of New South Wales, Sydney, NSW 2052, Australia (e-mails: {ihsanullah, r.malaney}@unsw.edu.au). L. Sun is with Nanjing University of Science and Technology (e-mail: sunlinlin@njjust.edu.cn). S. Tomasin is with the University of Padova, 35122 Padua, Italy (e-mail: tomasin@dei.unipd.it). The corresponding author is L. Sun.

low implementation cost. Furthermore, noting the recently emerging use of machine learning (ML) in the context of WVN, we summarize how ML can enhance the performance and be deployed in real-world LVS systems. Some critical aspects of ML-based LVS that can only be deployed with the aid of the B5G technology are also identified. Finally, future research directions and challenges in LVSs for WVN are discussed.

II. STATE-OF-THE-ART OF LOCATION VERIFICATION

In this section, we briefly highlight a few notable LVS frameworks post 2013, which may represent state-of-the-art of location verification in existing and future wireless networks.

A. Physical-Layer-Based Location Verification

Physical-layer-based LVSs utilize the inherent properties of the wireless medium to verify a user's reported location information and thus infer whether the user is legitimate or malicious. Their performance is based on the fundamental theories of detection, which serves as the performance limit of location verification in wireless networks.

In the LVS literature, the work [6], for the first time, developed an information-theoretic framework for location verification, proving that the likelihood ratio test is optimal in terms of maximizing the mutual information between the input and output of an LVS. This information-theoretic framework was deployed and examined in an LVS with received signal strength (RSS) measurements. The work [7] extended this framework by considering directional antennas under realistic dual-slope large-scale fading channel models in vehicular networks, showing that the directional antenna can significantly increase the capability of an LVS in correctly detecting malicious vehicles. Also in the context of vehicular networks, the authors of [8] considered location verification with multiple antennas under Rician fading channels. Interestingly, it was proved that the performance limit of LVSs is independent of the parameters of the channel from a malicious vehicle to a base station. In addition, as proved, the malicious vehicle's number of antennas does not affect this limit once it is greater than the legitimate vehicle's number of antennas. This is mainly due to the consideration of the worst-case scenario, where the malicious vehicle optimizes all the system parameters under its control. Because of this optimization, a further increase in the number of antennas held by a malicious vehicle has little additional impact.

LVSs also draw an increasing amount of research attention in the Internet of things (IoT). We note that WVN is one type of IoT, where the "things" are vehicles. For example, the work [9] developed an enhanced location verification using audibility and two-way time-of-arrival information in order to guarantee the reliability of location information for geo-spatial tagging and location-based services in IoT. Based on both synthetic and real-world datasets, it demonstrated that the LVS performance is improved when audibility is used. Considering the low-complexity hardware used to collect RSS measurements, they have been widely used in the context of location estimation and verification. In order to eliminate

the requirement of knowing transmit power in RSS-based systems, the differential RSS (DRSS) have been used as measurements as well. For the first time, [10] proved the identity of using RSS and DRSS in location verification for known and unknown transmit power, respectively.

B. Machine-Learning-Based Location Verification

Although LVSs based on information or statistical detection theory can offer performance limit of location verification, they require ideal operating conditions (e.g., known channel parameters), which may not be fully satisfied in practice. One motivation of using ML in practical LVSs is to eliminate this requirement. In this subsection, we briefly review some recent works on ML-based location verifications.

The authors of [11] developed an LVS based on neural network and demonstrated its efficient performance in the absence of prior knowledge on the proportion of genuine or malicious vehicles among all the vehicles. It was shown that the LVS based on neural network outperforms an information-theoretic LVS when the signals from the vehicles are under the impact of non-line-of-sight biases. The work [12] resorted to ML solutions for in-region location verification in order to solve the issues of lacking channel feature statistics. It was shown that the developed solution based on neural network and support vector machines, using typical loss functions, becomes the most powerful test at learning convergence for sufficiently complex learning machines and large training datasets, even in the absence of communication channel statistics. We note that the aforementioned typical loss functions refer to different loss functions used during the training phase of various ML algorithms, e.g., the cross-entropy loss function and the mean squared error loss function. Different from most aforementioned physical-layer-based LVSs that assume known channel models, the authors of [13] used logistic regression algorithm of the ML family for conducting location verification. This algorithm eliminates the assumption of known channel models or parameters and thus is applicable to more general scenarios of wireless networks. The performance of the proposed algorithm was examined based on RSS measurements at multiple landmarks (each is equipped with multiple antennas), where a distributed Frank-Wolfe-based verification is also developed in order to reduce the communication overhead. The detection accuracy of the proposed algorithm was confirmed by simulation and experimental results.

Based on the above existing works, we know that ML-based LVSs have the potentials of performing satisfactorily under variable conditions and in new environments. For example, an ML-LVS designed for an urban area may still function well in a suburban or rural area, as long as some training data is available. In Section IV, we will present more details to further demonstrate its high robustness.

III. SIMULTANEOUS LOCATION REPORTING AND VERIFICATION IN WIRELESS VEHICULAR NETWORKS

B5G wireless networks bring new opportunities to fully address the requirement of WVN. In this section, we identify a promising solution to location verification by using mmWave frequencies combined with multi-antenna techniques.

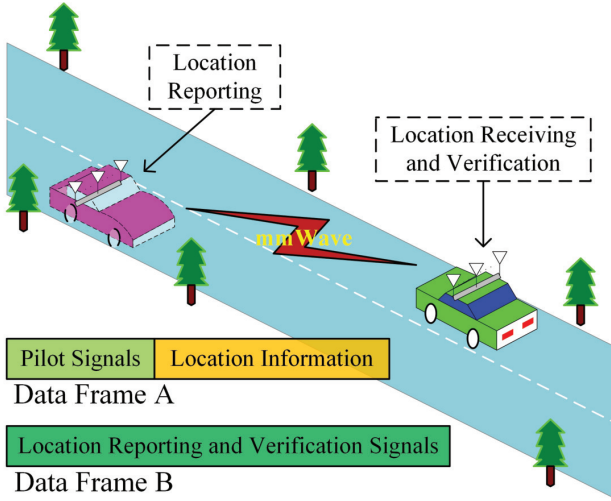


Fig. 1. The frameworks of location reporting and verification simultaneously with the aid of mmWave and multi-antenna techniques in B5G wireless vehicular networks.

A. Frameworks and Principles

An obvious benefit of enabling location reporting and verification simultaneously (LRVS) is the reduction of needed hardware resources and entailed processing delay. Communications via mmWave techniques can enhance LRVS in WVN mainly due to the high frequency (e.g., 28 GHz, 60 GHz) and high bandwidth (e.g., 400 MHz, 1 GHz) utilized. Also, the communication channels for mmWave are typically line-of-sight (LoS). Both these issues aid LRVS. The LRVS concept is inspired by emerging combined radar-communication techniques [14]. Although, the detection target and communication content are not closely related in radar-communication systems, in our proposed LRVS system the communication content is the location information of the detection target (e.g., a vehicle). Therefore, LRVS can be interpreted as an application of radar-communication, which demonstrates the benefits of joint communication and sensing. The multi-antenna techniques are useful for LRVS systems, since in many scenarios a vehicle may be connected to only one base station. In such a scenario a single-antenna system cannot guarantee the reliability of the reported location information. This is due to the fact that a malicious vehicle can modify its location metric (e.g. transmit power or transmission time) together with a false reported location in order to deceive an LVS [10].

We now present two frameworks of using mmWave combined with multi-antenna techniques to enable LRVS in WVN (see Fig. 1). The first framework involves the use of the data frame A shown in Fig. 1. This frame is of a form widely used for wireless communications based on the use of full CSI (coherent communications). In the context of LRVS the pilot signals of this frame are used for channel estimation and the location information part (indicated by the yellow shading in Fig. 1) is used simply to transfer the encoded GNSS coordinates. In addition to channel estimation, the pilot signals can also be used for location verification [8], since one benefit

of using data frame A for LRVS is a ‘double-dipping’ on the use of pilot signals - a technique that can significantly reduce latency.

A second framework for LRVS involves the use of the data frame B shown in Fig. 1. In this frame only statistical information of the channel is used for encoding (non-coherent communications). In this case a combined encoding of the claimed location coordinates and the verification information (e.g., signal strength and/or time-of-arrival) can be utilised. We note that the aforementioned two data frame structure does not require specific size, and in general a commonly used transmission control protocol (TCP) data size (e.g., 1500 bytes) is acceptable.

B. Challenges and Future Research Directions

Both these LRVS frameworks are different from previous studies of LVSs where the claimed location and the verification processes are considered to be separate and sequential phases of the LVS (e.g. data frame A used solely for conveying coordinates and then a separate set of signals sent for verification purposes). Which of the above frameworks for LRVS will prove to be more useful in B5G - enhanced WVN networks remains an open question. While data frame B may be more efficient, as it merges the two functionalities of data frame A in a single format, channel coding based on statistical CSI is not as efficient as that based on full CSI.

In general, communication and detection systems require different channel parameters for their own objectives. For example, in Rician fading channels for mmWave, an LVS requires the Rician parameter that determines the weight of the LoS component in these channels. Against this background, how to design transmit signals to simultaneously enable both the communication and detection systems is a main challenge in the context of LRVS - an issue further complicated by quantization errors and its impact on claimed location accuracy. Ultimately there exists a performance tradeoff between location reporting and location verification, which is highly affected by the resource (e.g., time slots and transmit power) allocation between the communication and detection systems. As such, a further challenge is how to optimally split resources between location reporting and location verification in order to achieve the best location verification. We note that ML may have the abilities to overcome these challenges. We discuss the use of ML in the context of location verification in the following section.

Data exchange among vehicles can also be observed by a fixed roadside network, that can independently verify the location of the vehicles (through either of the proposed data frames) and report its assessment to the cars. In this scenario the fixed network can also collect data points from many cars over time in its coverage area, thus obtaining a reliable training of ML solutions with a limited effort. Moreover, having multiple receive antennas at the roadside that sense the transmission of the same location-reporting vehicles strengthens the location verification and makes it more robust against sophisticated attacks by a vehicle using beamforming to the legitimate receiving vehicles.

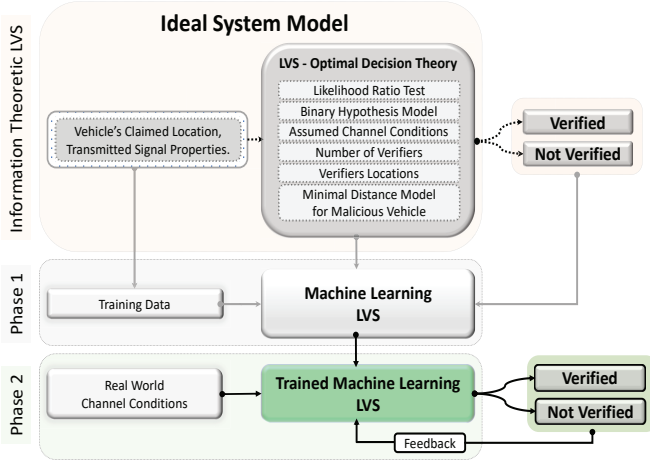


Fig. 2. A schematic of integrating machine learning (ML) with optimal decision theory to produce an LVS that is practically functioning in real-world channel conditions.

IV. MACHINE LEARNING FOR LOCATION VERIFICATION IN WIRELESS VEHICULAR NETWORKS

In this section, we first present some preliminary trials and conclusions of using ML in the context of location verification. Then, we present some principles and steps of using ML in solving location verification problems. Finally, we clarify and identify some challenges and future research directions on ML-based location verification for WVN.

A. Preliminary Trials and Results of Using ML in LVS

Many physical-layer-based LVSs reviewed in Section II will function optimally if the system model they assumed is a reality. However, this is nearly impossible in practice due to abrupt changes in channel conditions, changing noise parameters, diverse transceiver characteristics, and environment-based positioning errors. These issues may lead to a circumstance where an LVS designed for one application scenario may not function properly in other scenarios (e.g., with different channel conditions). For instance, an LVS designed for an urban environment may not function properly in a rural area (and vice versa). To accommodate all the above issues, we desire an LVS that can re-tune itself to the changing environments. The reviewed works mentioned in Section II confirm that ML plays a pivotal role towards the development of such an LVS within the context of WVN.

It is part of our own ongoing work to seamlessly integrate ML with optimal decision theory to develop an LVS solution for WVN that is practically deployable. A schematic of our approach to development of such a solution is shown in Fig. 2. We aim to integrate the information-theoretic LVS framework into neural networks. As shown in phase 1 of Fig. 2, we use the data considered for the information-theoretic LVS as training data for the LVS based on neural networks, where we aim to achieve similar location verification performance by neural networks to that of the information-theoretic LVS. In addition, the information-theoretic framework is able to provide us a benchmark performance that can aid the training of the neural

networks, e.g., to avoid the over training issues. In phase 2, we introduce real-world channel conditions to the ML-based LVS. Specifically, we plan to supply RSS, ToA, and AoA measurements of the vehicles' transmitted signals, and their claimed locations, as input to the LVS. Also, with the aid of existing and newly identified features in the input data and by adjusting the neural network architectures, we plan to further enhance the performance for the ML-based LVS. An active feedback from the verification output assists in continuously tuning the hyper parameters of the neural network, which leads to a scalable, adoptable, and up-to-date LVS in this context.

Some preliminary results achieved by our framework detailed above are presented in Fig. 3, where the total error is adopted as the performance metric. This total error is defined as $P_0\alpha + P_1\beta$, where P_0 and P_1 are the proportions of the legitimate and malicious vehicles, respectively, and α and β are the false positive and missed detection rates, respectively. To spoof the LVS, we consider that a malicious vehicle adopts an attacking strategy where it randomly claims its attack location at a predetermined distance (100 meters) away from its true location. We recall that, as discussed in Section III-A, the malicious vehicle not only reports a false location, but also intends to modify the measurements used by an LVS for verifying the reported location (in order to deceive the LVS). To plot Fig. 3, we assumed that the malicious vehicle has optimized its transmit power using the methodology presented in [10]. This is the worst-case scenario where the malicious vehicle optimizes all the parameters under its control.

An information-theoretic LVS [6], which is developed based on a likelihood ratio test (LRT), needs some *a priori* knowledge (e.g., P_1) for operation, while an LVS based on neural networks may not need such knowledge [11]. In the simulations, we train the neural network with incremental data as the time increases before testing the performance of the neural network by using the test data. Specifically, the number of training examples at 1 second is 1 and this number increases by 1 in each second. Therefore, we have severe fluctuations in the total error at the start, and polynomial fitting is used to smooth the total error curves. In this figure, we show that the neural network based LVS can still outperform the former, even when the neural network based LVS does not know P_1 (the values of P_1 shown in the legend of Fig. 3 are unknown). However, for the calculations in Fig. 3 it is assumed the LRT-based LVS knows the value of P_1 . In order to examine the robustness of the ML-based LVS, we consider different values of P_1 , ranging from very high (e.g., 50%) to very low (i.e., 0.05%). This result explicitly demonstrates the superiority of the ML-based LVS relative to the LRT-based LVS. We also see that the total error for the ML-based LVS converges once the framework is trained for approximately 800 seconds. The slight variation in the curves after 800 seconds is due to the applied polynomial fitting.

B. Principles and Steps of Using Neural Networks in LVS

Neural networks are chains of interconnected functions where one function processes the output of the one that precedes it. We take $f(x) = f^3(f^2(f^1(x)))$ as an example, where

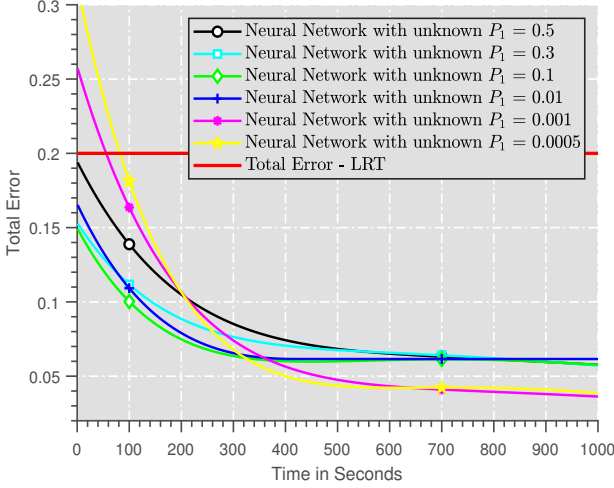


Fig. 3. Preliminary comparison results between an LVS based neural networks with unknown P_1 and an information-theoretic LVS based on likelihood ratio test (LRT) with known P_1 .

x is the input information. This input in our case includes the physical layer measurements and the vehicle's claimed location. In this example, f^1 , f^2 , and f^3 represent the first, second, and third layer of the neural network, respectively, where the final layer is the output layer as shown in Fig. 4.

How to determine a suitable neural network architecture for an LVS in WVN is an issue that can largely be resolved through extensive trials based on the available datasets. However, we note that determining an efficient neural network architecture is challenging due to many unknowns in the system. Here, we highlight a few key guidelines for facilitating the design of such a neural network architecture in the context of location verification for WVN.

- It is sufficient to start with a feedforward neural network with a single hidden layer and enough hidden units (neurons) to cover the anticipated number of unknowns.
- Increasing the number of layers and neurons (in each layer) can help in extracting the unknowns in the training data. However, we note that it is occasionally difficult to optimize a neural network based on such training data. During our intensive simulations with the available inputs, we observe that a single hidden-layer neural network may provide the best performance.
- If more features are identified as inputs in an LVS, a neural network with more depth and width can be potentially adopted. This can reduce the detection error probability achieved by the neural network (which is achieved based on the training data and the ground truth available information). However, we note that such deep neural networks may lead to over-fitting issues. In addition, more depth and width results in more computations and thus a trade-off exists between the desired performance and the corresponding complexity.
- It is suggested to choose transfer functions in the neurons that are close to linear, since they help in improving the LVS performance in WVN with acceptable complexities.

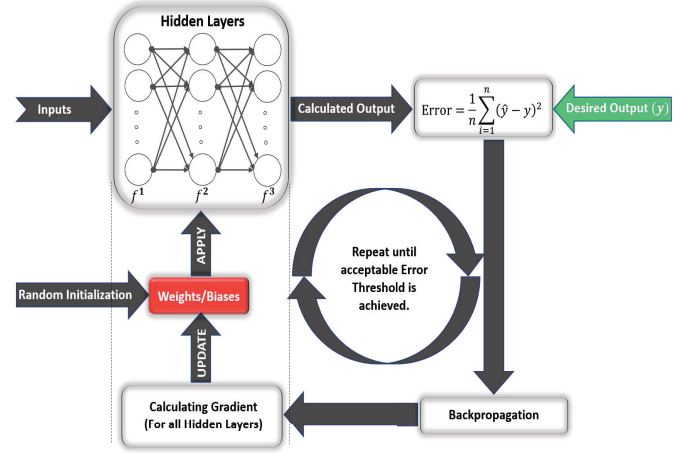


Fig. 4. Principles and steps of using neural networks in location verifications for B5G wireless vehicular networks.

Based on our comprehensive trials, we recommend the use of a linear rectified unit (ReLU) transfer function in the hidden layers and a pure linear transfer function in the output layer. The ReLU transfer function is half-linear and thus the complexity of the resultant optimization is relatively acceptable. On the other hand, the well-known logistic sigmoid and tangent sigmoid transfer functions saturate at absolute large values. As such, gradient-based learning is difficult with such functions. Thus, their use in the hidden layers is discouraged. To support this claim, we present Fig. 5 here, which is obtained from our detailed study of a neural network for LVS using different transfer functions. Two different spoofing scenarios from a malicious vehicle are shown in the figure. In both scenarios, the malicious vehicle randomly claims its location at some predetermined distance away from its true location. This distance is 50 meters in the first scenario (represented by the solid curves) and it is 75 meters in the second scenario (represented by the dashed curves). The curves in both scenarios relate to different transfer functions (refer to the legends in Fig. 5 for more details). In this figure, a better or equal performance in terms of achieving a lower or equal total error is obtained by the ReLU function (with a relatively lower complexity), compared to the logistic sigmoid and tangent sigmoid transfer functions.

- We suggest to determine a learning gradient for the neural network using a backpropagation algorithm. The weights and biases in different layers can then be updated through a learning process (e.g., by applying a stochastic gradient descent to the calculated gradients). A backpropagation algorithm that is fast at convergence should be adopted. Based on our extensive trials, we recommend the Levenberg-Marquardt backpropagation algorithm for using neural networks in location verifications for WVN.

C. Challenges and Future Research Directions

As an issue plaguing all ML methods, overfitting or over-training exists in the ML-based location verifications. This

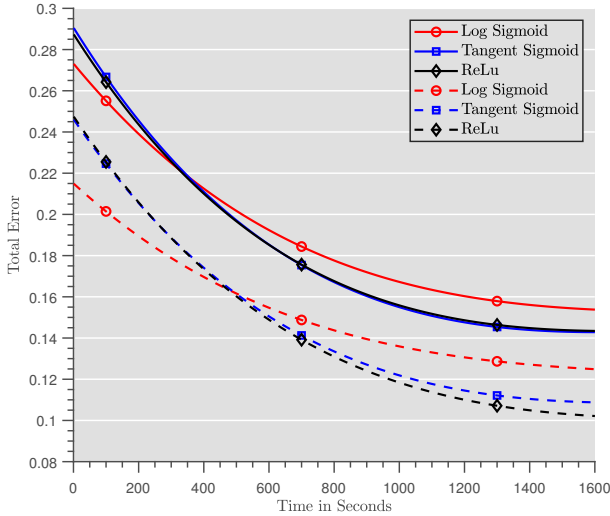


Fig. 5. Location verification performance of neural networks with different transfer functions in the hidden layers for B5G wireless vehicular networks.

is also the main reason to use simple and closely linear transfer functions in the neural networks for an LVS in the last subsection, as this issue is usually caused by a high model complexity. In addition, the overfitting issue becomes more visible when there exists more biases between the training data and real test data. In LVSs for WVN, these biases are due to many aspects, e.g., varying channel models, dynamic noise statistics, and uncertainties in threat models. Against this background, how to quantify these biases and then provide data support for avoiding overfitting issues is a challenging research problem in ML-based location verifications, which deserves a certain amount of research efforts.

As reviewed in Section II-B, ML methods have been increasingly used to address the unknown channel parameters in LVSs for WVN. In addition to these issues related to channel modelling, another major source of uncertainties in location verifications is the threat model, wherein the attacking strategies (e.g., the transmit power) and the true locations of the malicious user (e.g., the vehicle) should be specified. So far, most of the threat models for LVSs are developed by considering worst-case scenarios, where a malicious user's attacking strategy and true location are first optimized to minimize the detection performance of an LVS (e.g., [6]). Although these threat models can provide analytical results, they require some specific channel models or system parameters, which may not be available in practical application scenarios. In practical LVSs, we desire a threat model that is model-free and can update itself based on newly available data dynamically. Therefore, developing model-free, robust, and dynamic updating threat models is believed to be another advantage of using ML methods in the context of location verification for emerging WVN. The challenges in this context include (but are not limited to) the trustworthiness of all input data, the potential for overfitting, and identification of when to retrain the network updating.

A possible approach would be to make no assumptions

on the attacks, but only design an ML model to identify correct location reports - a problem that goes under the name of one-class classification. Neural networks used to perform one-class classification are called auto-encoders, and in [12] it was shown that an auto-encoder asymptotically (for large inputs and complex enough networks) performs as well the generalized likelihood ratio test. Within this context, we would be aiming at keeping under control false alarms (that do not depend on the attack model), while the missed-detection probability would be assessed by simulations or experiments with specific attack models.

Another promising direction to improve location verification is based on the integration (fusion) of many location-dependent services. For example, information partially dependent on location may come from other available reports within WVN such as broadcast messages, routing update messages, and radar reports. This additional information from various messages and reports, e.g., the relative location information among multiple vehicles (obtained based on the routing results as per the known protocols), can be transformed into new features and supplied as input to the ML-based location verification frameworks. Although each of these additional features alone would not be useful for localization or location verification, when analyzed collectively with the other existing input features, they may significantly improve location verification processes. Indeed, it has been previously shown how merging data from different sources makes user-authentication mechanisms more robust [15]. The ML solutions described in the previous sections can be adapted to such merging of multi-source data, by suitable training and adaptation. In addition, the emerging federated ML is another candidate for addressing location verification issues. An attractive feature of federated ML is that it preserves a user's privacy (e.g., keeps the location information private) while enabling the use of powerful ML techniques. It can also avoid the transmission of a large amount of raw data.

V. CONCLUSIONS

Verification of reported locations will be a critical function within emerging vehicular networks. In this work we have reviewed state-of-the-art techniques for delivering reliable location verification within real-world operational networks. Although optimal decision-making algorithms have been developed for many idealised channel conditions, we have argued that real-world verification solutions will most likely depend on ML-based algorithms, and have discussed initial research in this direction. Finally, we have peered into the future and presented some ideas on how emerging B5G solutions, combined with ML, could lead to even more robust forms of location verification.

REFERENCES

- [1] Y. Cao, T. Jiang, O. Kaiwartya, H. Sun, H. Zhou, and R. Wang, "Toward pre-empted ev charging recommendation through V2V-based reservation system," *IEEE Trans. Syst. Man Cybern.*, vol. 51, no. 5, pp. 3026–3039, May 2021.
- [2] G. Yan, S. Olariu, and M. Weigle, "Providing location security in vehicular ad hoc networks," *IEEE Wireless Commun.*, vol. 16, no. 6, pp. 48–55, Dec. 2009.

- [3] R. Lu, X. Lin, X. Liang, and X. Shen, "A dynamic privacy-preserving key management scheme for location-based services in VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 127–139, Mar. 2012.
- [4] Y. Cao, X. Zhang, B. Zhou, X. Duan, D. Tian, and X. Dai, "MEC intelligence driven electro-mobility management for battery switch service," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4016–4029, Jul. 2021.
- [5] J. Yang, Y. Chen, W. Trappe, and J. Cheng, *Pervasive Wireless Environments: Detecting and Localizing User Spoofing*. Springer, 2014.
- [6] S. Yan, R. Malaney, I. Nevat, and G. Peters, "Optimal information-theoretic wireless location verification," *IEEE Trans. Veh. Technol.*, vol. 63, no. 7, pp. 3410–3422, Jan. 2014.
- [7] M. Monteiro, J. Rebelatto, and R. Souza, "Information-theoretic location verification system with directional antennas for vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 93–103, Jan. 2016.
- [8] S. Yan, R. Malaney, I. Nevat, and G. Peters, "Location verification systems for vanets in Rician fading channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 4132–4144, Jul. 2016.
- [9] J. Y. Koh, I. Nevat, D. Leong, and W.-C. Wong, "Geo-spatial location spoofing detection for Internet of Things," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 971–978, Dec. 2016.
- [10] S. Yan, I. Nevat, G. Peters, and R. Malaney, "Location verification systems under spatially correlated shadowing," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4132–4144, Jul. 2016.
- [11] U. Ihsan, S. Yan, and R. Malaney, "Location verification for emerging wireless vehicular networks," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10 261–10 272, Dec. 2019.
- [12] A. Brighente, F. Formaggio, G. Nunzio, and S. Tomasin, "Machine learning for in-region location verification in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2490–2502, Nov. 2019.
- [13] L. Xiao, X. Wan, and Z. han, "PHY-layer authentication with multiple landmarks with reduced overhead," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1676–1687, Mar. 2018.
- [14] F. Liu, C. Masouros, A. Li, H. Sun, and L. Hanzo, "MU-MIMO communications with MIMO radar: From co-existence to joint transmission," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2755–2770, Apr. 2018.
- [15] H. Fang, X. Wang, and S. Tomasin, "Machine learning for intelligent authentication in 5G and beyond wireless networks," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 55–61, Oct. 2019.

Shihao Yan received his Ph.D degree from The University of New South Wales (UNSW), Sydney, Australia, in 2015. He was a Postdoctoral, University, and Senior Research Fellow at The Australian National University, Macquarie University, and UNSW, Australia, respectively. He is currently a Senior Lecturer in Edith Cowan University, Perth, Australia.

Ullah Ihsan received his B.Sc. degree from GIK Institute Pakistan. He pursued his PhD degree from UNSW Australia. He has been associated in technical roles with various wireless telecommunication operators and vendors for 12 years. Currently, he is the Lead Data Scientist with Xynoptik Australia in their digital transformation division.

Robert Malaney has a BSc. and Ph.D. from the University of Glasgow, Scotland, and the University of St. Andrews, Scotland, respectively. He is currently a Professor at the University of New South Wales, Australia. He was previously with the California Institute of Technology, USA, the University of California, Berkeley, USA, and the University of Toronto, Canada.

Linlin Sun received her B.S., M.S., and Ph.D. degrees from the Nanjing University of Science and Technology, Nanjing, China, in 2000, 2003, and 2020, respectively, where she is currently an Associate Professor with the School of Electronic and Optical Engineering.

Stefano Tomasin is an Associate Professor at the University of Padova, Italy. He has been on leave at Polytechnic University in Brooklyn and Huawei Research Laboratory in Paris. His research interests include physical layer security and signal processing for wireless communications, with application to cellular networks.