RBS-MLP: A Deep Learning based Rogue Base Station Detection Approach for 5G Mobile Networks

Mohammad Saedi, Adrian Moore, Philip Perry, Chunbo Luo, *Member, IEEE*

***Abstract*—The 3GPP Security Group has identified the detection of Rogue Base Stations (RBS) in 5G networks as one of the leading security challenges for users and network infrastructure. Motivated by this, RBS-MLP, a novel deep learning model, has been developed to identify RBSs. The model uses signal strength measurements in each mobile device's periodic measurement reports as input data, a reliable metric readily available to the system. We investigate the impacts of various sizes of datasets, different window sizes of received signal strength, and different proportional splits of the dataset into training and test data to evaluate the performance of the proposed model. We further demonstrate RBS-MLP using a realistic dataset of received signal strength measurements for a vehicle driving along various sections of a road, providing a use case to demonstrate the use of RBS-MLP to improve the safety of mobile networks. Experimental results reveal that RBS-MLP is well suited as a 99.999% accuracy classification model and provides a new baseline method for RBS detection.**

***Index Terms*—Rogue Base Station (RBS), 5G Mobile Networks, Attack Detection, Vehicle Platooning, Machine Learning (ML), Received Signal Strength (RSS), Measurement Report (MR), gNodeB.**

I. Introduction

ROGUE Base Stations (RBS) are wireless devices that impersonate a legitimate Base Station (BS), causing subscribers within a certain radius to connect to those devices rather than genuine networks. An RBS attack can happen during the initial cell search stage in the 5G NR when a User Equipment (UE) looks for a suitable BS to camp on. During this stage, the UE listens to the wireless broadcast channel for the synchronising signal (SS) from nearby BSs. Next, the UE selects a BS based on the received SSs and initiates a wireless connection. If an RBS broadcasts a spoofing SS with high Received Signal Strength (RSS) during the cell search mechanism, the UE may be enticed to it and try to camp on it rather than any legitimate BS [1]. In the emerging 5G world, it will be vital for infrastructure providers to protect against such attacks to secure the communications platform and protect client data and identity.

Since the inception of early GSM networks, RBS attacks have continuously evolved and persisted. These can be categorised as Denial of Service (DoS) attacks on mobile devices or networks, provision of fraudulent services, and compromising of subscribers' privacy. The impact of these attacks varies greatly among cellular network generations but remains significant owing to the multiple interconnections across a diverse set of current and legacy networks [2]. The development of 5G communications has already led to some advancements in RBS detection, such as Subscription Permanent Identifier (SUPI) concealment, guaranteed Globally Unique Temporary Identity (GUTI) refreshment, and protected redirections. At the same time, other security mechanisms inherited from previous generations include mutual authentication between UE and network, secure algorithm negotiations, and integrity-protected signalling [3]. Despite these advances, the current position is that 5G remains vulnerable to RBS attack [4].

Most current RBS detection systems implement a data-gathering capability in the UE, which then either (i) performs analysis on the data gathered at UE-side, or (ii) sends the collected data to a central server for cloud-side detection, or (iii) to the wider network for analysis known as network-based detection [5]–[8]. Of these, the first group is prone to false positives because a UE cannot understand the complete status of the network view at any given time, and in addition, UE-side detection systems often need software updates on the device or root privileges, which are uncommon and may be difficult for some users. In the second group, the devices send their Measurement Reports (MR) to a central server in the 5G Core for analysis; otherwise, they operate on the same premise as UE-side detectors and, as a result, suffer from the same efficacy and scalability concerns. The third group, Network-based detection systems, are projected to perform better in terms of analysis since, unlike UEs, mobile networks have knowledge of the system's global status. However, a supporting monitoring infrastructure to collect data from multiple network locations or protocols is required. We focus on the third group, which is more applicable to 5G and cellular systems.

In recent years, there has been a lot of interest in Machine Learning (ML) for detecting security threats. ML models can identify security assaults by learning the behaviour of both attack and legitimate scenarios. ML-based classification systems can provide high levels of precision in the identification of potential aggressors [9], [10]. We use supervised classification methods to detect attacks in 5G vehicular platooning [11]. This paper presents the design, implementation, and development of an RBS detection system named RBS-MLP, exploring ML methods to detect RBS in 5G networks effectively. The proposed approach is essentially based on MR filtering. We present a case study demonstrating RBS detection in a realistic dataset of radio information and Received Signal Strength (RSS) measurements generated by a simulation of a vehicle travelling along various road sections.

M.Saedi is with the Department of Computing, Sheffield Hallam University, Sheffield, S1 2NU, United Kingdom (e-mail: m.saedi@shu.ac.uk)

A. Moore, and P. Perry are with the School of Computing at the Ulster University, Belfast, United Kingdom, (e-mail: {aa.moore,p.perry}@ulster.ac.uk)

C.Luo is with the Department of Computer Science, University of Exeter, Exeter, EX4 4RN, United Kingdom (e-mail: c.luo@exeter.ac.uk)

RBS-MLP is aimed to protect IoT devices by filtering RBS from MR to detect all rogue agents with as few false positives as possible, without specialised hardware. In summary, this research study provides the following contributions:

- Providing a 3GPP Release 18 compliant RBS detection model. The device-assisted part of the model uses the standard measurement reporting procedure, while the network-based part performs the data analysis using a deep learning RBS detection system.
- Evaluating the performance of the RBS-MLP algorithm in terms of the accuracy of classification between rogue and legitimate BS.
- Proposing an enhanced handover protocol to include a BS trust mechanism to evaluate the trustworthiness of the BS.
- Using the ML approach in a simulation of a platoon moving along sections of roads containing a mix of legitimate and rogue base stations.

The rest of the paper is structured as follows. In Section II, we review the related research in RBS detection systems while identifying their limitations. Section III presents a novel RBS detection model comprising a pre-processing component and a decision maker. Section IV proposes the RBS-MLP model. Section V presents the performance evaluation and results for various case studies based on our realistic dataset of radio information and RSS measurements taken by a simulated vehicle travelling along various sections of a road. Finally, the conclusions and future works are discussed in Section VI.

II. Related Work

In this section, we report a summary of existing RBS detection systems, mentioned in the previous section, according to the techniques used. In addition, drawbacks and some limitations will be described here.

The UE-side includes client-side applications that perform the identification within the UEs. This includes mobile phones, vehicles, IoT devices, etc. Android IMSI-Catcher Detector (AIMSICD) [5], Cell Spy Catcher [12], CatcherCatcher [13], and SnoopSnitch [6] are some applications that fall into this group. To provide some level of protection, these apps require high privileges and low-level access to baseband chips to reach their full potential. Even though Cell Spy Catcher and AIMSICD results have not been persuasive, Cell Spy Catcher at least can be used to determine if the local network figures have been modified. SnoopSnitch seems to be the most advanced of the alternatives, as it reliably informs the user immediately after the threat is detected. In contrast, Cell Spy Catcher only provides a warning and associated information. SnoopSnitch, on the other hand, only works on Qualcomm-based Android phones and requires root access. Similarly, CatcherCatcher attempts to detect RBS activity by detecting irregularities in mobile networks, but it only works on Osmocom phones. To summarise, these apps are still in the early stages of development for detecting RBS attacks. They have a lower detection rate, generate more false positives, and require unusually high-level access, making them unsuitable for the general public [14].

Techniques for cloud-based detection are based on analysing the crowdsourced data from a nearby massive number of UEs to detect and geolocate RBS units. FBS-Radar [7], a large-scale RBS detection and localisation system, identifies an RBS through the automated collection of suspicious SMS messages from end-user devices. In addition, these reports, including Received Signal Strength (RSS), cell identifier and UE MAC addresses, are sent to a server to analyse and evaluate different techniques that exploit this data to identify RBS installations accurately without analysing the content of the SMS messages. Van Do et al. [15] suggested a methodology for detecting abnormal behaviour from an RBS in public data sets using ML approaches. In [16], the experiment was extended using machine learning and exploiting a signature-based strategy with characteristics such as location and the relationship between the identification number of the UE and subscription. These investigations used a publicly available data set from Aftenposten [17] to demonstrate the utility of ML approaches, but with the drawback that UEs must report their measurements to a server on the cloud for analysis; otherwise, they operate like a client-side detector. As a result, they suffer from effectiveness and scalability issues.

Network-based detection techniques conduct the analysis on the core of the cellular network. In [8], a technique for IMSI catcher detection has been proposed that uses existing operational data from the mobile network used in the mobility management of mobile stations. The MRs delivered by the UEs to BSs containing information on the cell and surrounding cells are used to detect IMSI catchers. Regarding analysis, network-based detection systems are projected to outperform client-side detectors since mobile networks know the system's global status, unlike UEs. However, a limitation with [8] is that they only cover 2G radio access technology. Murat [2] is a network-based approach for recognising RBSs on several 3GPP Radio Access Technologies (RAT) without changing mobile phones or monitoring equipment. Murat employs the global state to include information on all connected mobile phones, the mobile network state, and its deployment and setup history. It outperforms earlier systems.

Traditional RBS techniques are purposefully planned, making it difficult for networks to adjust dynamically. ML is the process of self-learning from experiences and deeds and such approaches can be used to respond appropriately in such situations without human intervention or reprogramming. Deep learning, a subset of machine learning, has grown in popularity in recent years and has been used for RBS detection [18]; studies have shown that deep learning outperforms previous approaches [19].

III. Proposed Approach

In this section, first, the framework of the detection model is proposed. Then a novel handover protocol will provide a key feature for the ML method to build the main component of the RBS detection system.

The analyser component of the RBS-MLP is an ML-based approach based on MR; it can be considered an additional feature for Murat [2]. Murat offers a network-based approach for identifying bogus base stations that run on any 3GPP radio access technology without needing to modify mobile phones.

However, the analyser of Murat consists of data processors and Rule-based methods; we analyse the MR using ML methods [20], [21].

A. Proposed Architecture

Fig. 1 describes the proposed architecture to model an RBS identification system to detect rogue agents using machine learning methods. The UE in RRC_CONNECTED mode builds the MR based on signals received from gNBs currently in range and sends it to the 5G Radio Access Network (RAN). The proposed system performs the data analysis, identifies suspected RBS and eliminates them from consideration for handover [22]. As a result, the suspected RBS is never included in the version of the MR used to assess the need for a handover event. The BS identifies the need for a handover and, if required, initiates the protocol. Upon detection of an RBS, the network operators can be informed so that legal action and other post-incident activities can be initiated. For example, they can alarm the UE from camping on the RBS [23], [24].

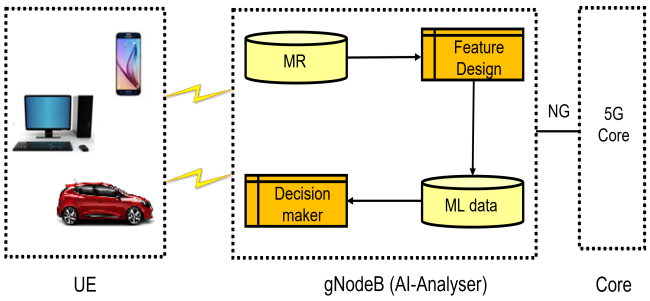


Fig. 1. RBS detector.

Such analysers might be included in either the 5G RAN or the 5G Core, however in this work, the integration of the analysis into the 5G RAN is examined. This decision can be explained by the need for scalability, which is most readily done when the RBS detector is situated at the point where the gNodeB receives the MR data [20].

Fig. 1 illustrates the components of the AI Analyser, comprising a Feature Design element, an AI-based Decision-maker element, and two MR and ML data storage units. The aim of Feature Design is to create informative and relevant input features that help the model distinguish between the two classes. The ML approaches will be used in the decision-making function. The Decision Maker, as the significant component of the AI Analyser, will receive the ML data and apply the ML method to identify rogues. Other Analyser functions can employ various strategies for classification, but we will demonstrate the effectiveness of an ML approach. The sections that follow expand on the architecture's description and details.

B. Handover Process

In a 5G environment, an IoT device typically has a selection of BS units within its reception range and the way the device decides which BS to attach is based on an analysis of the device's MR. The MR identifies the most vital received BS signals from the current location and orientation and is updated periodically, usually every second. Depending on the received signal strength of the currently connected BS and the signal strengths of alternate candidate BS units, the device will be told by the network either to stay with the current BS or handover to a more robust alternative. Handover management is critical for ensuring that UEs may move freely between cells while still receiving high-quality communication services. The gNB is in charge of managing UE migration across cells. Typically the handover decision in 5G RAN is based on the MR produced by UEs [25]. The handover process is that the UE is currently connected to a gNB (the serving BS); if there is another BS in the MR with a level that exceeds the threshold value, then we handover to it. Fig. 2 illustrates the Received Signal Strength (RSS) from a collection of 30 BS encountered by a platoon leader along a stretch of urban motorway. As the UE moves along the highway, the handover process is activated, resulting in the connected signal strength profile illustrated in Fig. 3.

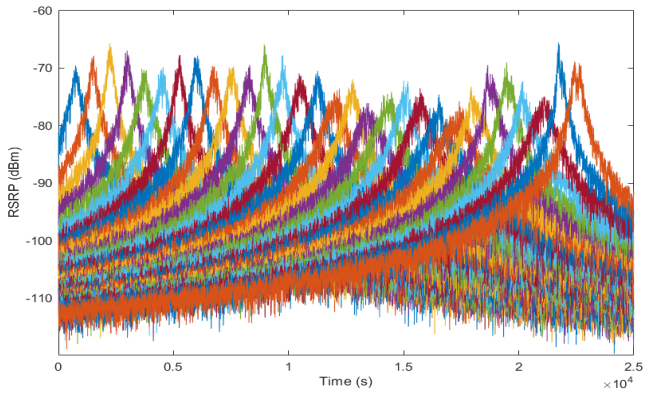


Fig. 2. Received signal strength in a legitimate scenario.

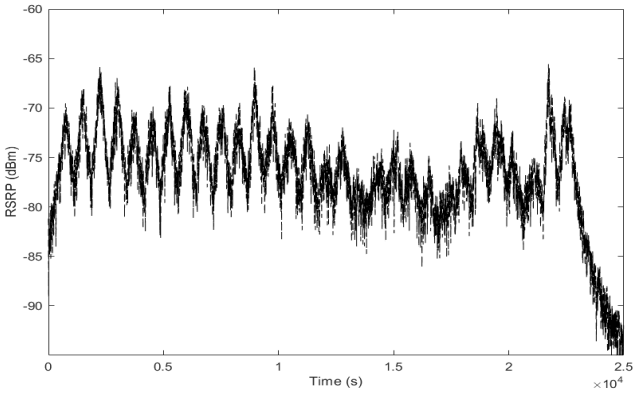


Fig. 3. Handover decision-making in a legitimate scenario.

C. Feature Design

To entice surrounding UEs, an RBS often transmits higher-than-normal signal power. A steep increase and fall in RSS is a characteristic of the profile of a typical rogue [11], which can

be used as a fingerprint for detection. Therefore, this section defines and implements two new features including the Rate of Change (RoCH) and Probation Period. These informative features are calculated and monitored for all BS and RBS that UE detects. The main idea is that the RoCH is a better indicator for RBS detection than "raw" RSS data, as RBS tend to increase at a more rapid rate than "regular" BS. Following is a description of the Probation period.

Fig. 4, for instance, demonstrates BS1, the connected BS, with decreasing power over time as BS2 rises. The RoCH of BS2 can be collected if its readings during the Probation Period are monitored. The Probation Period begins when the BS is first detected and concludes when sufficient consecutive MR values have been recorded. The optimal number of values will be determined later in the following sections. In the diagram, the timestamp "C" indicates the beginning of the Probation Period or Candidate BS Monitoring Period, and the timestamp "H" indicates the transition time.

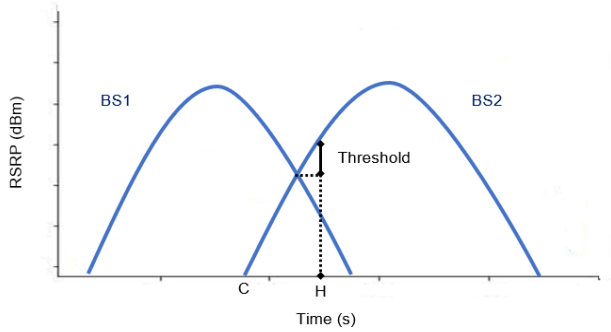


Fig. 4. Candidate BS monitoring period for a legitimate scenario.

The analysis is a continuous procedure that proceeds regardless of any handover. The RBS (in red) in Fig. 5 is recognized as a rogue before the probable handover point. The next sections will describe various possibilities depending on whether both BS and rogue could be in the MR at the same time or not. However, a rogue agent must not be taken into account for handover.

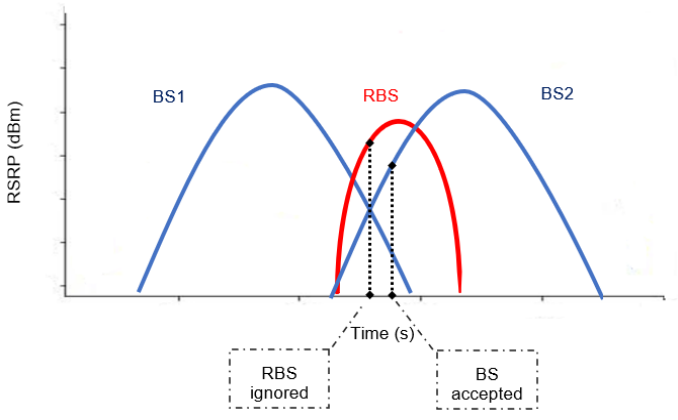


Fig. 5. Detection of rogue base stations and transfer to a legal base station.

The BS Analysis system will use RoCH as one of its essential features to learn any fingerprint related to the RBS. In addition, it will be considered in the ML approach by constructing a fingerprint of the rogue data stream and attempting to identify what is in a dataset that characterises it as legitimate or rogue.

D. Proposed Analysis System

By analyzing the RSS value ranges, it can be observed that the MR's strongest signal values are around -70 and -75, while the weakest signals are around -90 and -95. These values are influenced by the distance between the BS and the road and the power of the BS transmission. The graph in Fig. 6 shows an attack scenario where RBS values increase and decrease more rapidly compared to LBS values. Consequently, the average rate of RoCH for RBS is significantly higher during the initial rise period than for LBS. As the RoCH is dependent on the speed of the vehicle, the system needs to learn and adjust the threshold accordingly.

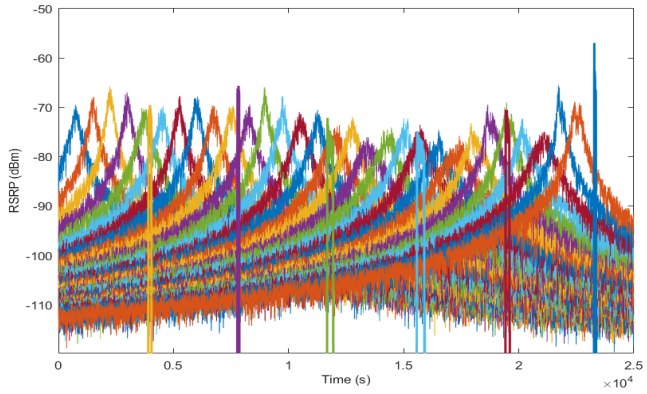


Fig. 6. Attack scenario.

Fig. 7 demonstrates the transfer to RBS in this particular scenario. The timeframes when the platoon leader is linked to an LBS are denoted by a blue signal, while the connection to an RBS is indicated by a red signal.

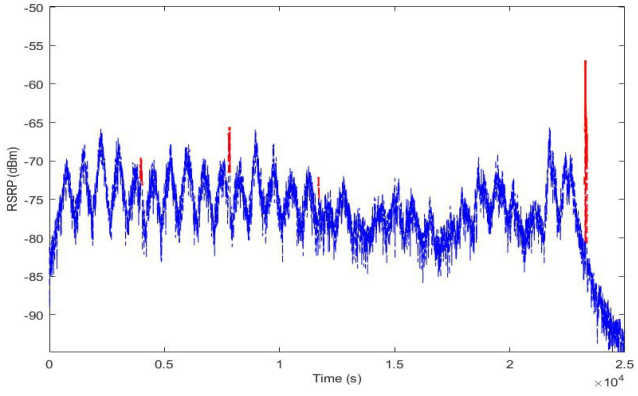


Fig. 7. Handover decision-making in an attack scenario.

This research introduces a novel BS monitoring model that takes into account the potential presence of rogue actors in the

mobile network. According to this proposal, which is depicted in Fig. 8, a state machine is designed with three states for each BS, including a blocked state that is reserved for a BS that has been identified as a rogue in the MR and should not be considered for handover:

- Blocked state: a BS that will not be considered for handover. This may be because it has been assessed as potentially rogue by the classifier, or because it has been recently discovered and not yet completed its probation period.
- Candidate state: a BS that has completed its probation period and has not exhibited any rogue-like characteristics. If the RSS of a candidate BS exceeds that of the currently connected BS, then a handover event is initiated.
- Connected state: the currently "active" BS. Only one BS will be in this state at any time.

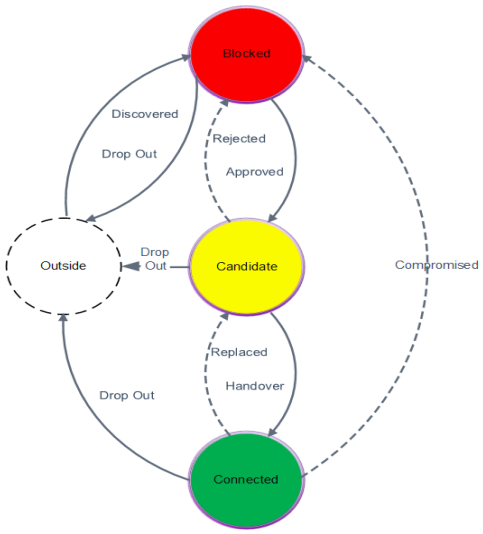


Fig. 8. BS state machine in Analyser.

A newly discovered BS should be blocked until the probation period has expired. Upon completion of the probationary period, the BS will be approved. Each BS moves between the states depending on the rate of change of its signal strength. For example, a candidate exhibits a sudden large jump in its RSS value resulting in it being classified as a rogue and becoming blocked (Rejected). Alternatively, if a candidate rises normally and becomes the strongest signal exceeding the handover threshold, then it will be the subject of a handover event and will become connected (handover). If the currently connected BS is no longer the strongest, it will be replaced and revert to the candidate state (replaced), or the connected BS might exhibit rogue-like behaviour so would be immediately blocked. The Drop Out arc from other states to the "Outside" state machine represents a BS that drops out of the MR. In reality, it needs to move out of range and then be rediscovered and pass a probation period to become a candidate. A "Discovered" BS cannot become a candidate until it has been observed and checked for a complete window of timestamps. During this period, it is in the Blocked state. The aim is to remove the arcs connecting "Candidate" and "Connected" with "Blocked" by detecting RBS before they are considered for detection.

IV. Deep Learning Approach

This section of the study intends to investigate the effectiveness of the ML method in distinguishing between rogue and legitimate signals by training it with both types of instances. The aim is to eliminate the need for a predetermined threshold level in the process.

A. Produce Measurement Report

The 3GPP Technical Report 38.331 [26] outlines UE measurement reports that contain pertinent data to identify Rogue Base Stations (RBS). The reports include the cell identity and Received Signal Strength (RSS), as well as information about cell groups (CGI_info) that incorporate data from both the MIB and SIB1.

The simulation tool utilised in the research can simulate a vast number of BS and RBS along an extensive motorway, as discussed in [27]. However, not all data related to BS and RBS can be accessed by the handover mechanism simultaneously. According to the 3GPP specification, the MR component stores only the six strongest Received Signal Strength (RSS) values at any given time [28], [29], which are calculated based on the received signals from the six most prominent BS within the platoon leader's vicinity, as depicted in Fig. 9.

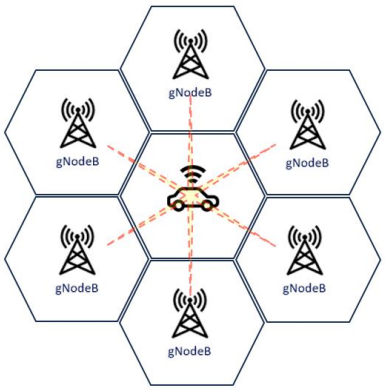


Fig. 9. Multiple data streams represent BSs, but only a maximum of six BSs is included in MR at any time.

B. Machine Learning Dataset

When a new BS is discovered, it is assigned the "Blocked" state. Once it has been present in the MR for a defined number of consecutive samples, it can be analysed to determine whether it is legitimate ("Candidate") or rogue ("Blocked"). Therefore, it is the first run of consecutive values that should be used as training data for the proposed model and the optimal length of this initial run is determined in the following sections.

The Feature Design component in Fig. 1 takes the accumulated MR data over a simulation period and generates the ML dataset which will be utilised specifically for training the

TABLE I
DATA FILE OF ML DATA (L: LEGITIMATE, R: ROGUE)

BS(i)	Timestamp(i)										BS Target
	RSS1	RSS2	RSS3	RSS4	RSS5	RSS6	RSS7	RSS8	RSS9	RSS10	
BS(1)	-76.601	-76.545	-76.426	-75.997	-75.659	-75.418	-75.334	-75.154	-75.040	-75.583	L
BS(2)	-82.753	-81.928	-81.835	-81.686	-80.905	-82.026	-82.161	-81.919	-81.874	-81.666	L
BS(3)	-84.585	-84.817	-85.761	-85.991	-86.332	-86.516	-86.462	-86.439	-85.777	-85.655	L
BS(4)	-87.612	-88.044	-87.889	-86.993	-87.592	-87.623	-88.213	-87.058	-87.102	-87.290	L
BS(5)	-90.285	-89.245	-89.290	-89.188	-88.509	-87.579	-87.804	-88.252	-87.731	-87.887	L
BS(6)	-90.862	-90.232	-90.018	-89.431	-89.628	-88.791	-90.088	-89.592	-89.791	-89.952	L
BS(7)	-89.820	-89.463	-89.647	-89.154	-69.260	-68.773	-65.956	-66.554	-65.613	-66.044	L
BS(i-1)	-79.483	-74.951	-69.992	-69.564	-69.679	-70.132	-69.947	-69.899	-71.056	-70.889	R
BS(i)	-83.950	-80.313	-75.188	-70.416	-70.399	-70.267	-70.876	-70.851	-70.891	-71.973	R
BS(n-1)	-84.039	-79.276	-75.123	-75.509	-74.913	-75.110	-75.424	-75.446	-75.697	-76.000	R
BS(n)	-92.281	-87.595	-83.476	-78.500	-74.190	-74.293	-74.622	-74.371	-74.691	-75.132	R

ML classifier. In addition, the rate of changes can be added as well to have more features for learning more to identify RBS accurately.

The experiments carried out in this study necessitate using a dataset that contains both malicious and legitimate BSs. The simulations include different road lengths and BS/RBS positions/densities [11]. When the UE (in this scenario, the lead vehicle of a platoon) is within the range of the BSs, the received signal is calculated per second.

A snapshot of the ML training data, which includes data streams from both LBS and RBS, is shown in Table I. Each BS is assigned its own line in the data, with the first set of consecutive RSS readings for that BS in the MR and an identification (L for LBS, R for RBS) following.

The width of the sample window is the quantity of RSS samples included in each BS set. Alternative window sizes will be looked into even though the size of the window in this example is assumed to be 10.

Three datasets of different sizes have been created as described in Table II. For example, in the first dataset, we simulate a 500 km road with 30000 timestamps in which there are 90 legitimate BSs and 18 rogue BSs. The details of the three datasets are presented.

TABLE II
PARAMETER SETTING OF SIMULATIONS OF THREE DATASETS

Dataset	LBS	RBS	Road Length(Km)	Timestamp(S)
90LBS–18RBS	90	18	500	30000
500LBS–90RBS	500	90	4000	170000
1000LBS–180RBS	1000	180	5000	225000

C. Classification

The classification stage in Machine Learning is a procedure for determining whether or not an observation falls into a specific category. Here, it detects whether an unknown base station (observation) is genuine or fake (categories).

In this section, it will be demonstrated how datasets and features were combined to generate classifier models to correctly identify a stream of received signal values as representing either a legitimate or rogue BS. The classifier is trained by feeding it successive data streams (one stream at a time), indicating for each whether the stream represents a legitimate or rogue BS.

Once the classifier has learned how to differentiate between legitimate and rogue streams, we can then pass it an unknown stream representing the output from either a legitimate or rogue BS and have it classified.

An artificial neural network algorithm is implemented to verify the model performance. The detection model proposed is a binary classification Multilayer Perceptron (MLP) using the sequential API as shown in Fig. 10. The classifier includes three hidden layers. The first layer consists of “relu” activation function with a ‘he_normal‘ weight initialisation to overcome the problem of vanishing gradients when training deep neural network models. The activation functions in the second layer are the “tanh” and “sigmoid” functions. The optimiser is “SGD” (Stochastic Gradient Descent), and the loss function is “binary_crossentropy”. The selection of all activation functions and loss functions was based on the analysis results.

V. PERFORMANCE EVALUATION AND RESULTS

In this stage, we will apply our classifier to test data sets to evaluate the accuracy and reliability of our method. First, we will consider some evaluation metrics in the following paragraph and then show some diagrams to show and compare the results with the following scenarios.

- True positive (TP): positive samples correctly classified as positive, here, i.e., correctly identified LBS. This would remain in the "Candidate" or "Connected" state in the state machine provided in Section III.
- True negative (TN): negative samples correctly classified as negative, i.e., correctly identified RBS; refer to a state diagram in which BS would move into (or remain in) the “Blocked” state.
- False-positive (FP): negative samples incorrectly classified as positive, here, i.e., wrongly identified LBS; BS which should be “blocked” remains as “candidate”. When an RBS has been identified as a legitimate (FP), it makes a critical situation in the handover process – a situation that we are trying to prevent.
- False-negative (FN): positive samples correctly classified as negative, i.e., wrongly identified RBS; In this case, the BS which is a legitimate “candidate” is wrongly

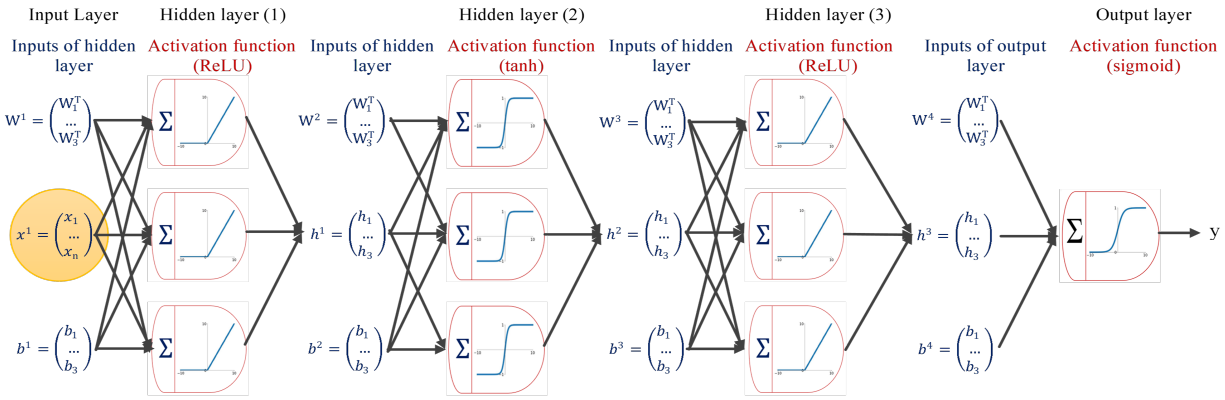


Fig. 10. Artificial neural network architecture.

"blocked". Not the desired outcome, but not a disaster as long as we are currently connected, or another legitimate candidate is available. Therefore, when a legitimate BS is identified as a rogue, the impact is less significant.

Four different metrics are used for each classifier's evaluation [30], [31]. First, accuracy is the classifier's ability to categorise the samples:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Second, Recall or Sensitivity, is the proportion of positive samples that are classed as positive. It is also called Sensitivity or True Positive Rate (TPR), which is the LBS detection probability:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

Third, precision is the proportion of correct positive classifications (TP) from cases predicted as positive:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Fourth, the F1-score is the harmonic mean and takes precision and recall into account:

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

Next, we examine the model's performance and investigate the effects of different data set sizes, RSS window sizes, and different portions of training and test data sources. Through experimentation, we compared three datasets with 70/30 splits between training and testing as well as varied window widths of RSS in terms of accuracy, recall, precision, and F1-score. As might be predicted, accuracy may be observed to grow as more data is taken into account. A 70/30 split between training and testing data results in an accuracy result of 0.975 for WS=3 with 500 LBS and 90 RBS, as indicated in Fig. 11 by a dashed green line.

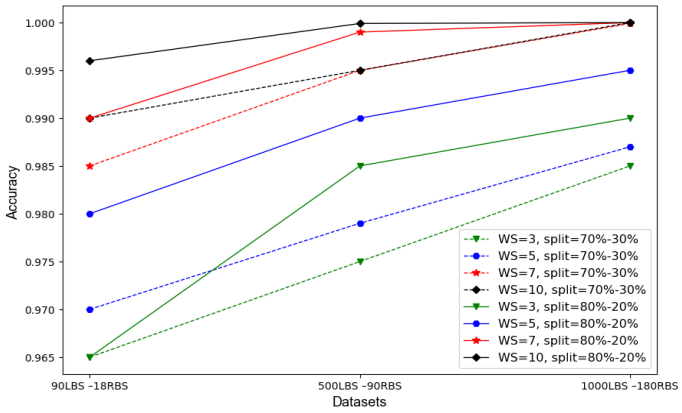


Fig. 11. Accuracy metric.

The green line indicates that this metric increases to 0.985 with an 80/20 split between training and testing data. Accuracy rises even further for the bigger dataset, hitting 0.995 for the 500LBS-90RBS dataset and 0.99999 for the 1000LBS-180RBS dataset. Similarly, the precision, recall, and F1-score metrics will rise with larger datasets and training data, which will be explored in the following part.

The greater the window size, the more data there will be to work with to make a more accurate conclusion. The Accuracy measure in Fig. 11 shows that WS=3 is not a sufficiently trustworthy size, and so it is not an adequate size for a window. WS=5 is substantially more reliable, although WS=7 yields 0.995 accuracy. However, for the biggest dataset, WS=10 is the best choice overall, with accuracy=0.99999. However, when it comes to the overhead of using a bigger window size than is necessary, it is more likely that a larger window may hamper connection since a possible BS will be blocked for a longer amount of time, which may result in latency. Blocking a BS from handover for a longer length of time owing to a wider window may have a detrimental influence on signal availability to the platoon.

The decision-making process becomes more accurate as the range of possibilities widens. Nevertheless, even after exploring WS=12 and WS=15, it was found that they did not yield better results than WS=10. Consequently, the experiment

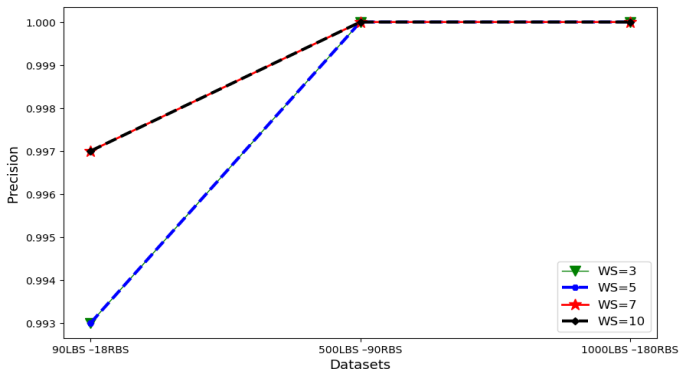


Fig. 12. Precision metric.

was terminated at WS=10, and now WS=10 can be fully trusted. There is no justification to consider larger window lengths at this point. Fig. 12, Fig. 13 and Fig. 14 show the findings for the other three performance metrics, revealing that the precision, F1-score, and recall factors for the 90LBS-18RBS dataset are 0.997, 0.998, and 1.0, respectively, and improve with larger comprehensive datasets.

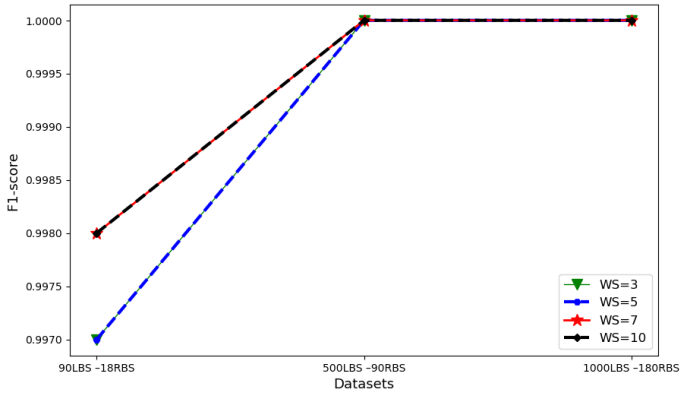


Fig. 13. F1-score metric.

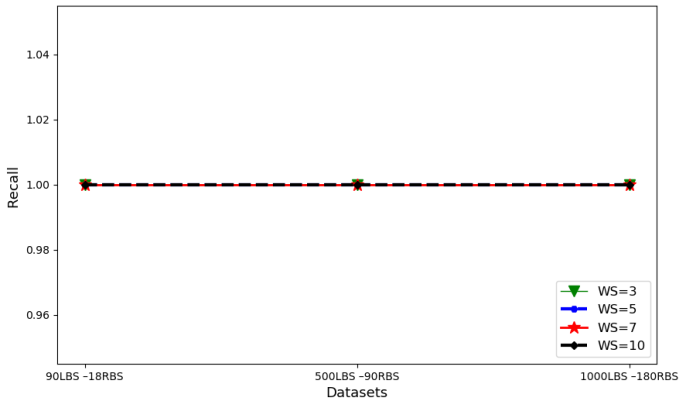


Fig. 14. Recall metric.

Table III demonstrates that the True Negative Rate (TNR), varies between 98.97% and 100% for various datasets with varying divisions of training and test data, implying that the detection probability of RBS for the proposed model is about 99.50%. On the other hand, the maximum bound of the False Positive Rate (FPR) that detects rogue agents as genuine is around 1%. The 500LBS-90RBS dataset appears to be an anomaly, outperforming the 1000LBS-180RBS dataset. Moreover, the False Negative Rate (FNR), which measures the probability that LBS is a rogue, is also zero.

In general, when the dataset is much smaller, there is less training data available, resulting in FP. Except for 500LBS-90RBS, all datasets attain an FPR close to zero. Plotting the loss function rate is a helpful technique to see if the model is appropriately trained. During model training, the loss function is utilised to determine the target value for the model to achieve. We have used binary cross-entropy from the Keras library, a high-level neural network library, in our experiment. It is used in binary classification model as a loss function and computes the difference in cross-entropy between true and predicted labels. This is crucial to ensuring that the model is fitted correctly. Fig. 15 shows the loss rate of the model for the same configuration. The minimal rate is accomplished in a variety of datasets. The results reveal that the loss rate is unaffected by window size; nevertheless, the larger the dataset, the lower the loss rate [32], [33].

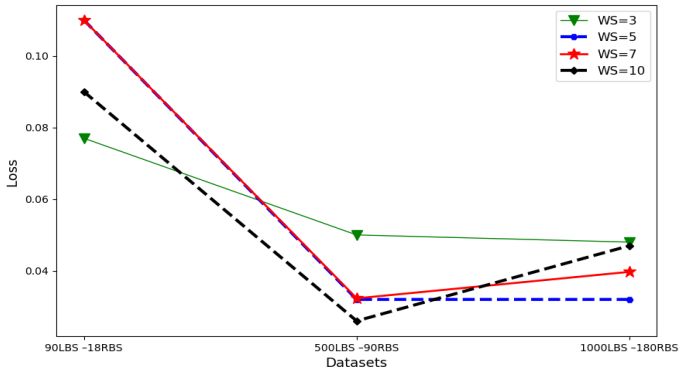


Fig. 15. Loss rate.

VI. CONCLUSION AND FUTURE WORK

5G includes many measures to help secure the network and protect users' privacy and security. However, one of the most pressing issues in user and network security, as recognised by the 3GPP Security Group, is the identification of malicious agents in mobile networks. RBS attacks have been identified as a significant threat. So, the detection of RBS attacks will be a substantial contribution to knowledge from this research. We have designed and implemented a new method to intelligently detect the presence of a rogue by using a Multi-Layer Perceptron technique. The system is termed RBS-MLP and it is 3GPP compliant. The system can be deployed in the gNBs of a 5G RAN to gather information from Measurement Reports from mobile devices and analyse how the reported signal strength varies with time to identify a signal from an RBS. We have tested this system with large sets of synthetic data from a vehicle platooning scenario to fine tune the system. The results show that the ML approach produces 99.999% accuracy, i.e., is one misclassification in

TABLE III
TPR: LBS DETECTION PROBABILITY, TNR: RBS DETECTION PROBABILITY

	FPR				FNR			
Dataset	WS=3	WS=5	WS=7	WS=10	WS=3	WS=5	WS=7	WS=10
90LBS–18RBS	0	0	0	0	0	0	0	0
500LBS–90RBS	1.03	1.03	1.03	1.03	0	0	0	0
1000LBS–180RBS	0	0	0	0	0	0	0	0

	TPR				TNR			
Dataset	WS=3	WS=5	WS=7	WS=10	WS=3	WS=5	WS=7	WS=10
90LBS–18RBS	100	100	100	100	100	100	100	100
500LBS–90RBS	100	100	100	100	98.97	98.97	98.97	98.97
1000LBS–180RBS	100	100	100	100	100	100	100	100

every 100,000, with some specific results and provides a new baseline method for RBS detection. Also, will help to improve 5G services in our target deployment (platooning). The next step of the work will involve the deployment of various AI/ML techniques and compare performance metrics.

ACKNOWLEDGMENTS

This research has been supported by the BT Ireland Innovation Centre (BTIIC) project, funded by BT, and Invest Northern Ireland.

REFERENCES

[1] X. Li, A. Garcia-Saavedra, X. Costa-Perez, C. J. Bernardos, C. Guimarães, K. Antevski, J. Mangues-Bafalluy, J. Baranda, E. Zeydan, D. Corujo *et al.*, "5growth: An end-to-end service platform for automated deployment and management of vertical services over 5g networks," *IEEE Communications Magazine*, vol. 59, no. 3, pp. 84–90, 2021.

[2] P. K. Nakarmi, M. A. Ersoy, E. U. Soykan, and K. Norrman, "Murat: Multi-rat false base station detector," *arXiv preprint arXiv:2102.08780*, 2021.

[3] H. Alrashede and R. A. Shaikh, "IMSI Catcher Detection Method for Cellular Networks," *2nd International Conference on Computer Applications and Information Security, ICCAIS 2019*, pp. 1–6, 2019.

[4] 3GPP Technical Report 33.809, "Study on 5G Security Enhancement against False Base Stations (FBS)," *3rd Generation Partnership Project*, no. Release 17, 2023.

[5] AIMSICD. Android imsi-catcher detector. Accessed on 2022-11-20. [Online]. Available: https://github.com/CellularPrivacy/Android-IMSI-Catcher-Detector

[6] SRLabs. Snoopsnitch, imsi catcher score. Accessed on 2023-01-05. [Online]. Available: https://opensource.srlabs.de/projects/snoopsnitch/wiki/IMSI_Catcher_Score

[7] Z. Li, W. Wang, C. Wilson, J. Chen, C. Qian, T. Jung, L. Zhang, K. Liu, X. Li, and Y. Liu, "Fbs-radar: Uncovering fake base stations at scale in the wild," *Internet Society*, 2017.

[8] S. Steig, A. Aarnes, T. Van Do, and H. T. Nguyen, "A network based imsi catcher detection," in *2016 6th International Conference on IT Convergence and Security (ICITCS)*. IEEE, 2016, pp. 1–6.

[9] D. Chulerttiyawong and A. Jamalipour, "Sybil attack detection in internet of flying things-ioft: A machine learning approach," *IEEE Internet of Things Journal*, vol. 10, no. 14, pp. 12 854–12 866, 2023.

[10] M. R. Dey, M. Patra, and P. Mishra, "Efficient detection and localization of dos attacks in heterogeneous vehicular networks," *IEEE Transactions on Vehicular Technology*, 2023.

[11] M. Saedi, A. Moore, and P. Perry, "Synthetic generation of realistic signal strength data to enable 5g rogue base station investigation in vehicular platooning," *Applied Sciences*, vol. 12, no. 24, p. 12516, 2022.

[12] Cell Spy Catcher (Anti Spy). Skibapps. Accessed on 2022-15-12. [Online]. Available: https://play.google.com/store/apps/details?id=com.skibapps.cellspycatcher

[13] CatcherCatcher. Mobile network assessment tools. Accessed on 2023-01-03. [Online]. Available: https://opensource.srlabs.de/projects/mobile-network-assessment-tools/wiki/CatcherCatcher

[14] B. Brenninkmeijer, "Catching imsi-catcher-catchers: An effectiveness review of imsi-catcher-catcher applications," *Bachelor Thesis, Radboud University (Nijmegen, The Netherlands)*, 2016.

[15] T. Van Do, H. T. Nguyen, N. Momchil, and V. T. Do, "Detecting imsi-catcher using soft computing," in *Soft Computing in Data Science: First International Conference, SCDS 2015, Putrajaya, Malaysia, September 2-3, 2015, Proceedings 1*. Springer, 2015, pp. 129–140.

[16] V. T. Do, P. Engelstad, B. Feng, and T. van Do, "Strengthening mobile network security using machine learning," in *Mobile Web and Intelligent Information Systems: 13th International Conference, MobiWIS 2016, Vienna, Austria, August 22-24, 2016, Proceedings 13*. Springer, 2016, pp. 173–183.

[17] Aftenposten. (2015) Aftenposten public data set. Accessed on 2023-01-05. [Online]. Available: https://www.aftenposten.no/meninger/kommentar/i/9mrn5/derfor-publiserer-aftenposten-hele-datagrunnlaget-for-mobilspionasje-s

[18] Z. Dong, K. Kane, and L. J. Camp, "Detection of rogue certificates from trusted certificate authorities using deep neural networks," *ACM Transactions on Privacy and Security (TOPS)*, vol. 19, no. 2, pp. 1–31, 2016.

[19] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21 954–21 961, 2017.

[20] P. K. Nakarmi and K. Norrman, "Detecting false base stations in mobile networks," p. 5, 2018. [Online]. Available: https://www.ericsson.com/en/blog/2018/6/detecting-false-base-stations-in-mobile-networks

[21] 3GPP TS 33.501, "Security architecture and procedures for 5G system," *3rd Generation Partnership Project*, no. Release 17, 2022.

[22] 3GPP TR 38.300, "NR and NG-RAN Overall Description," *3rd Generation Partnership Project*, no. Release 16, 2021.

[23] U. Gorrepati, P. Zavarsky, and R. Ruhl, "Privacy protection in lte and 5g networks," in *2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC)*. IEEE, 2021, pp. 382–387.

[24] Component, D. H. S. and Callahan, "Privacy Impact Assessment, (U.S. Department of Homeland Security," 2019. [Online]. Available: https://www.dhs.gov/xlibrary/assets/privacy/privacy{_}pia{_}template.pdf

[25] A. Shaik, R. Borgaonkar, S. Park, and J.-P. Seifert, "On the impact of rogue base stations in 4g/lte self organizing networks," in *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, 2018, pp. 75–86.

[26] G. T. S. 38.331, "Radio resource control (rrc) protocol specification," *3rd Generation Partnership Project*, vol. 0, 2021.

[27] M. Saedi, A. Moore, P. Perry, M. Shojafar, H. Ullah, J. Synnott, R. Brown, and I. Herwono, "Generation of realistic signal strength measurements for a 5g rogue base station attack scenario," in *2020 IEEE Conference on Communications and Network Security (CNS)*, June 2020, pp. 1–7.

[28] D. van Thanhe, I. Jørstad, and D. van Thuan, "Strong authentication for web services with mobile universal identity," in *Mobile Web and Intelligent Information Systems: 12th International Conference, MobiWis 2015, Rome, Italy, August 24-26, 2015, Proceedings 12*. Springer, 2015, pp. 27–36.

[29] R. Barco, F. J. Cañete, L. Diez, R. Ferrer, and V. Wille, "Analysis of mobile measurement-based interference matrices in gsm networks," *IEEE Vehicular Technology Conference*, vol. 3, pp. 1412–1416, 2001.

[30] V. Selis and A. Marshall, "A classification-based algorithm to detect forged embedded machines in iot environments," *IEEE Systems Journal*, vol. 13, no. 1, pp. 389–399, 2018.

[31] S. Gyawali, Y. Qian, and R. Q. Hu, "Machine learning and reputation based misbehavior detection in vehicular communication networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8871–8885, 2020.
[32] G. Ian, B. Yoshua, and C. Aaron, "Deep learning: Adaptive computation and machine learning," 2017.
[33] R. Reed and R. J. MarksII, *Neural smithing: supervised learning in feedforward artificial neural networks*. Mit Press, 1999.

To,

Prof Abbas Jamalipour,
The Editor-in-Chief,
IEEE Transactions on Vehicular Technology,

We are pleased to submit our research article "**RBS-MLP: A Deep Learning based Rogue Base Station Detection Approach for 5G Mobile Networks**" for consideration for publication in your esteemed journal.

Novelty/Original Contribution:

1. RBS-MLP is a novel deep learning and a 3GPP-compliant model to identify Rogue Base Station (RBS).
2. RBS-MLP uses a realistic dataset of received signal strength measurements for a vehicle driving along various sections of a road, providing a use case to demonstrate the use of RBS-MLP to improve the safety of mobile networks.
3. RBS-MLP Proposes an enhanced handover protocol to include a BS trust mechanism to evaluate the trustworthiness of the BS.
4. RBS-MLP Uses the ML approach in a simulation of a platoon moving along sections of roads containing a mix of legitimate and rogue base stations.
5. RBS-MLP produces 99.999% accuracy; provides a new baseline method for RBS detection. Also, it will help to improve 5G services in Vehicular Platooning.

Original Article Statement:

a. This manuscript is the authors' original work and has not been published nor has it been submitted simultaneously elsewhere.

b. All authors have checked the manuscript and have agreed to the submission.

We hope you find our manuscript suitable for publication and look forward to your favourable response.

Dr Mohammad Saedi
Senior Lecturer in Computing
Department of Smart Computing
Sheffield Hallam University
Sheffield, S1 1WB, UK
Tel: +44 (0)1142255000 (int. n. 9102)
Web: Mohammad | Sheffield Hallam University (shu.ac.uk)
E-mail: m.saedi@shu.ac.uk
Reviewer: IEEE Transaction on Vehicular Technology, IEEE Systems Journal, IEEE Access; member: IEEE Communications & IEEE Vehicular Technology Society

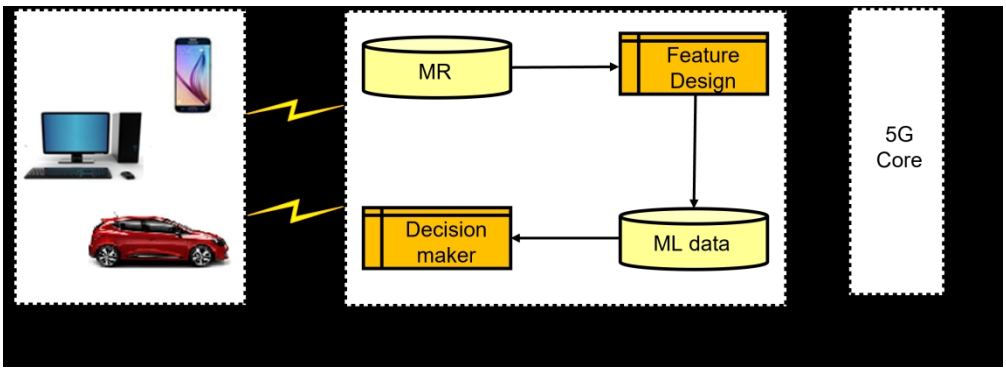


Fig. 1. RBS detector.

586x211mm (130 x 130 DPI)

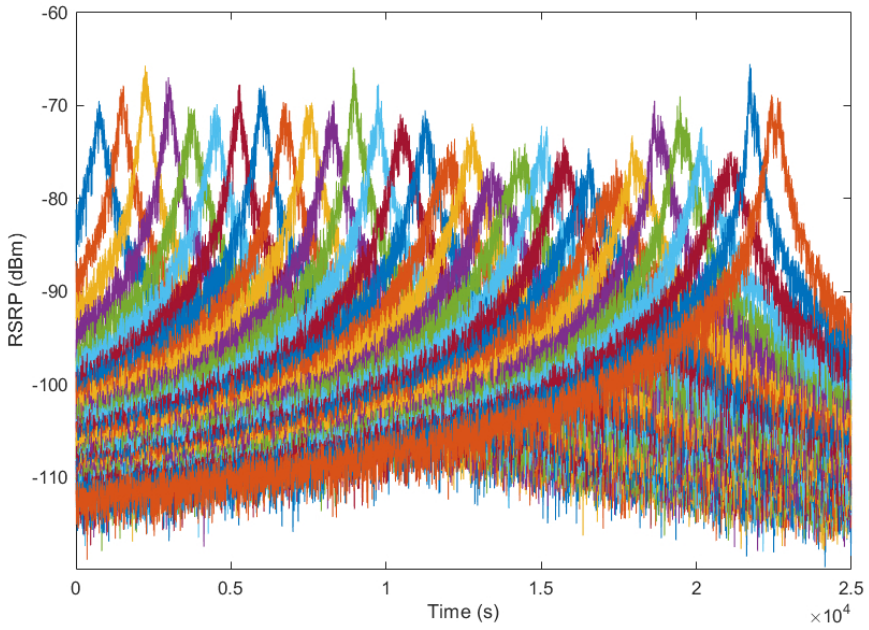


Fig. 2. Received signal strength in a legitimate scenario.

622x425mm (38 x 38 DPI)

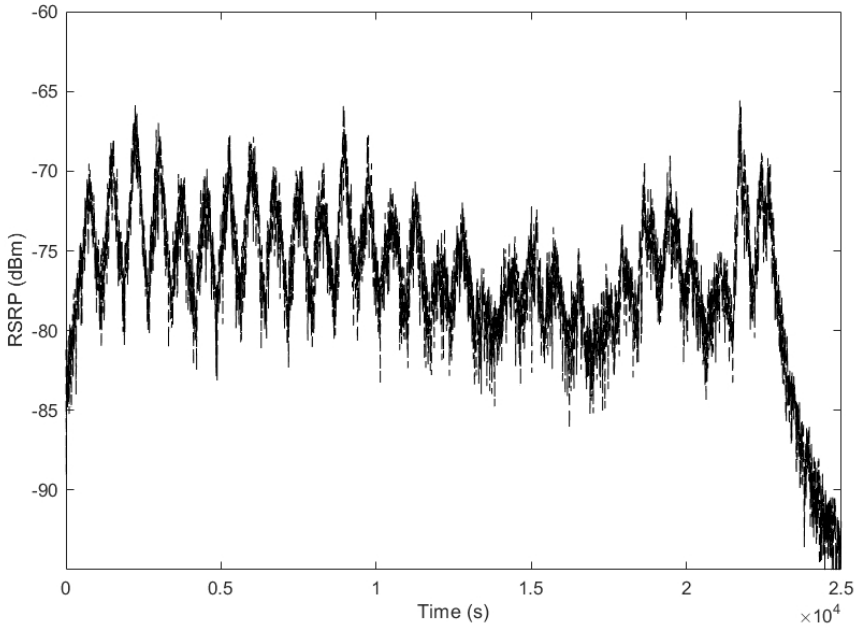


Fig. 3. Handover decision-making in a legitimate scenario.

622x425mm (38 x 38 DPI)

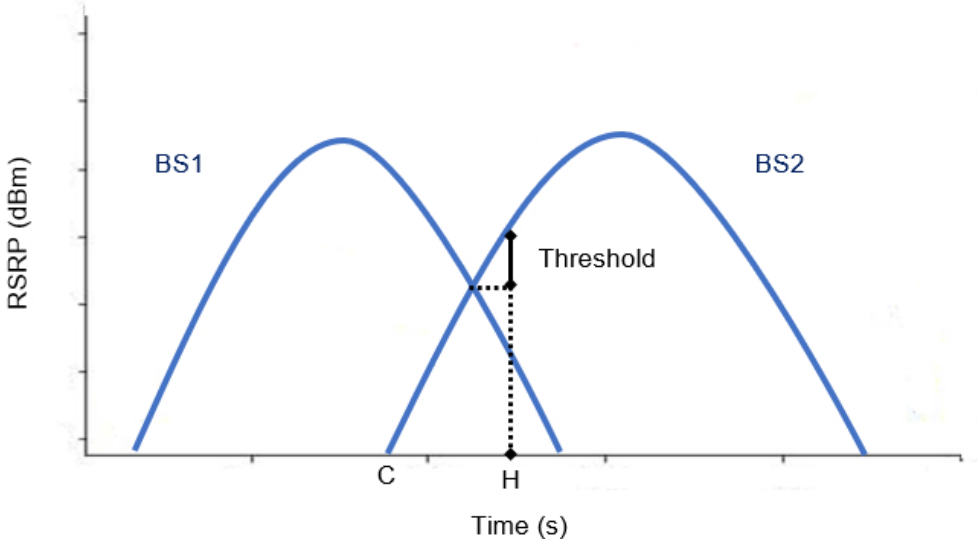
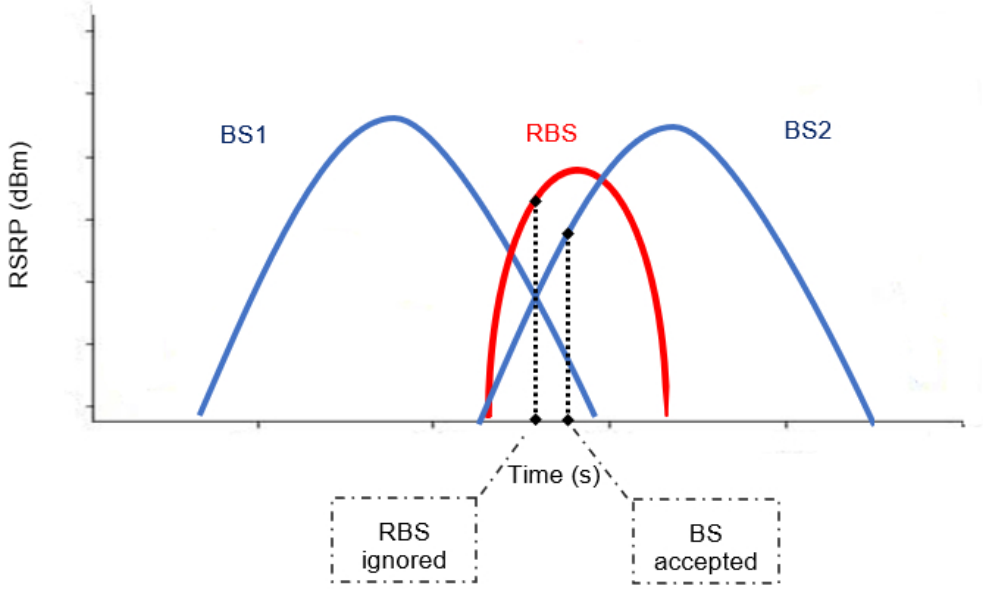


Fig. 4. Candidate BS monitoring period for a legitimate scenario.

354x199mm (47 x 47 DPI)



Detection of rogue base stations and transfer to a legal base station.

366x219mm (47 x 47 DPI)

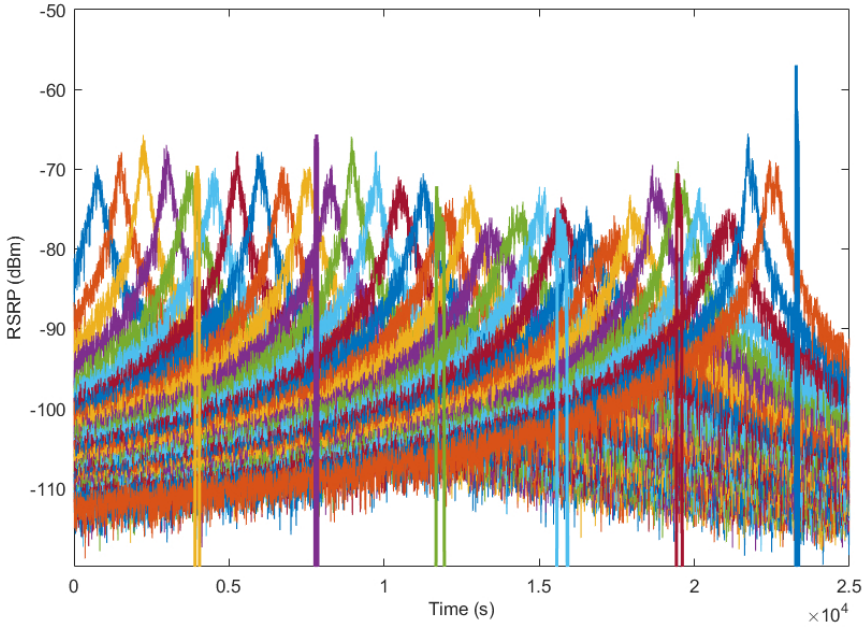


Fig. 6. Attack scenario.

622x425mm (38 x 38 DPI)

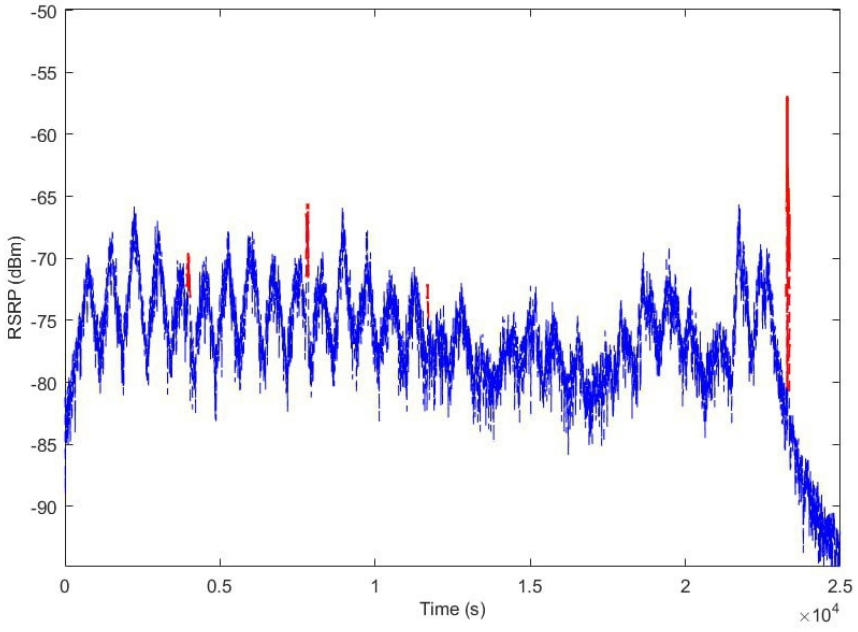


Fig. 7. Handover decision-making in an attack scenario.

328x224mm (72 x 72 DPI)

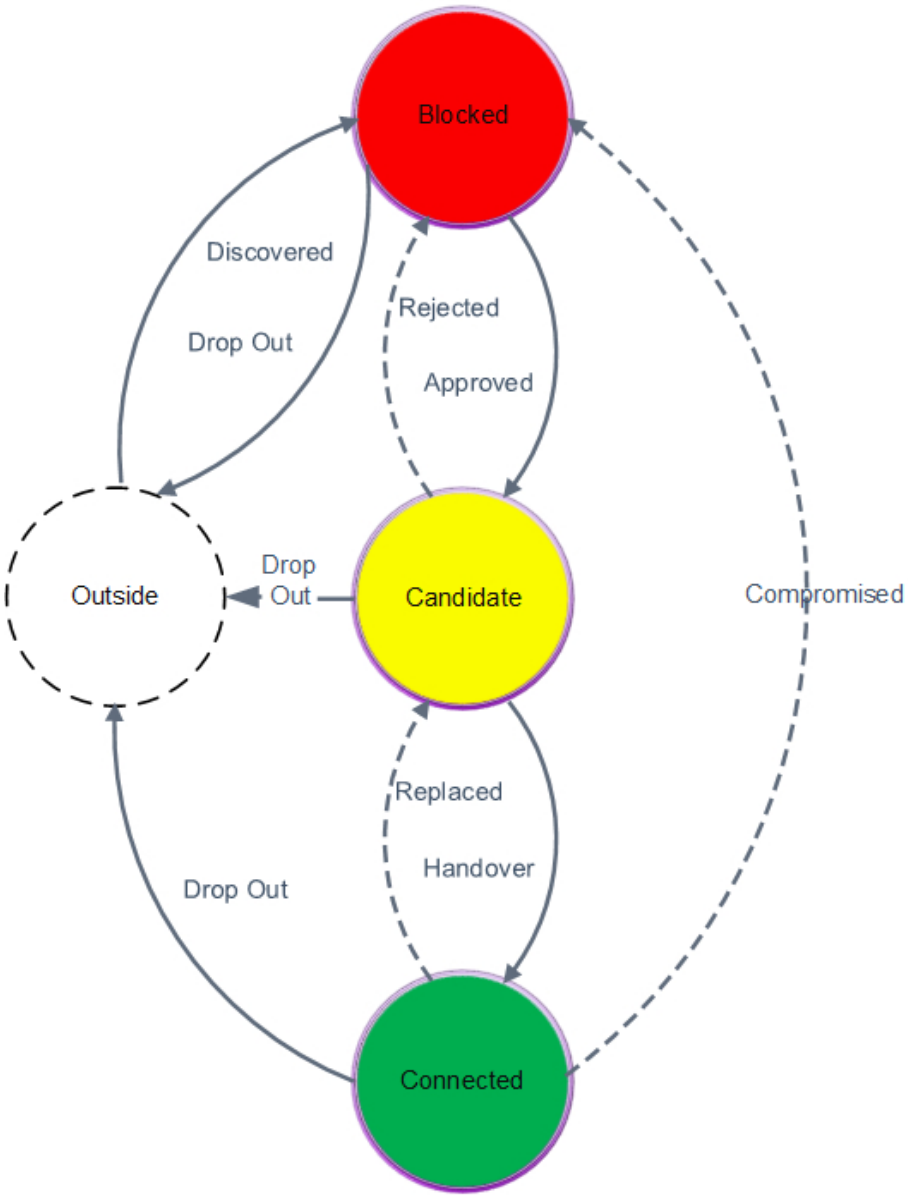


Fig. 8. BS state machine in Analyser.

288x377mm (47 x 47 DPI)

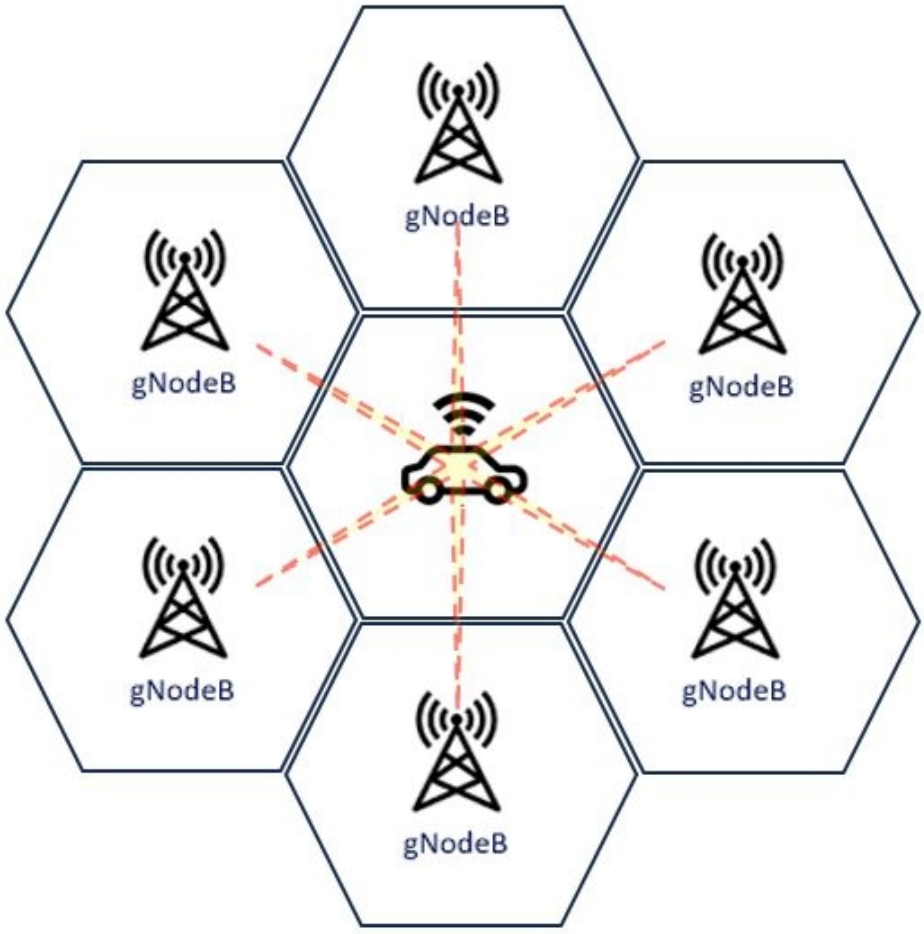


Fig. 9. Multiple data streams represent BSs, but only a maximum of six BSs is included in MR at any time.

117x111mm (120 x 120 DPI)

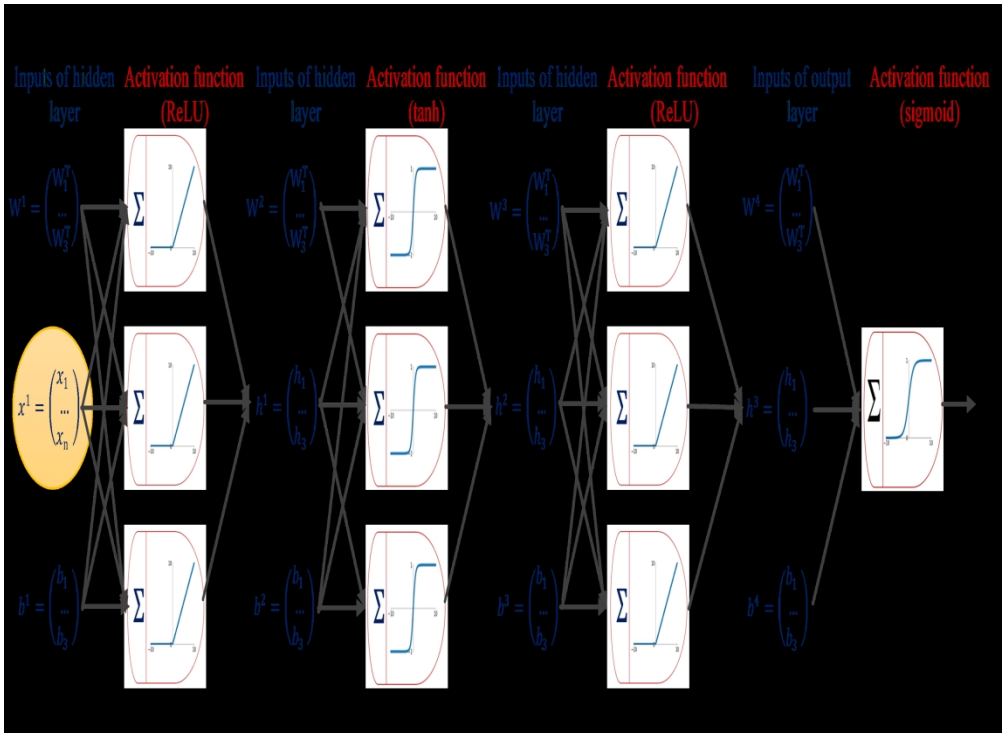


Fig. 10. Artificial neural network architecture.

370x270mm (236 x 236 DPI)

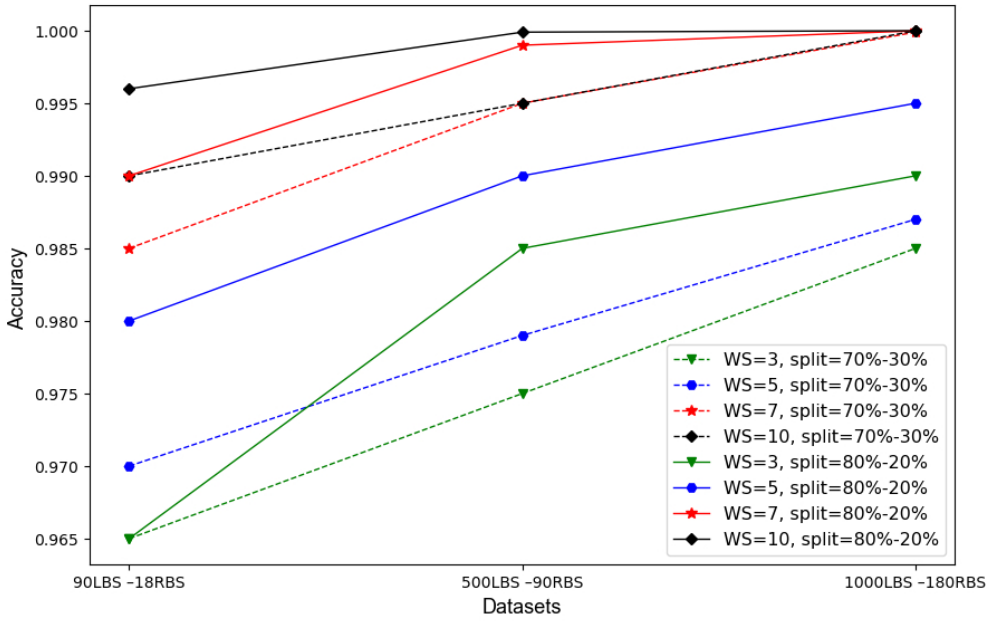


Fig. 11. Accuracy metric.

583x369mm (39 x 39 DPI)

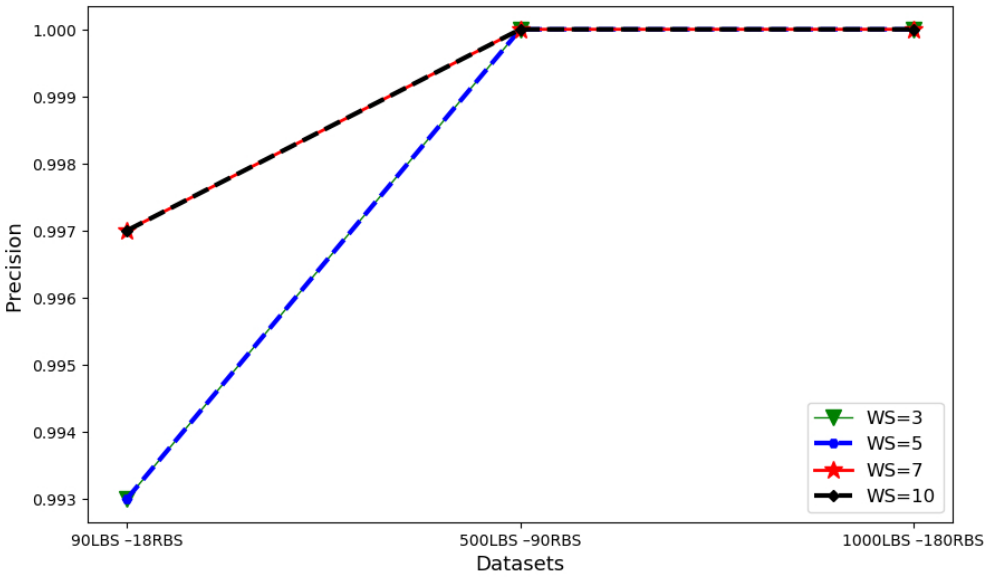


Fig. 12. Precision metric.

583x344mm (39 x 39 DPI)

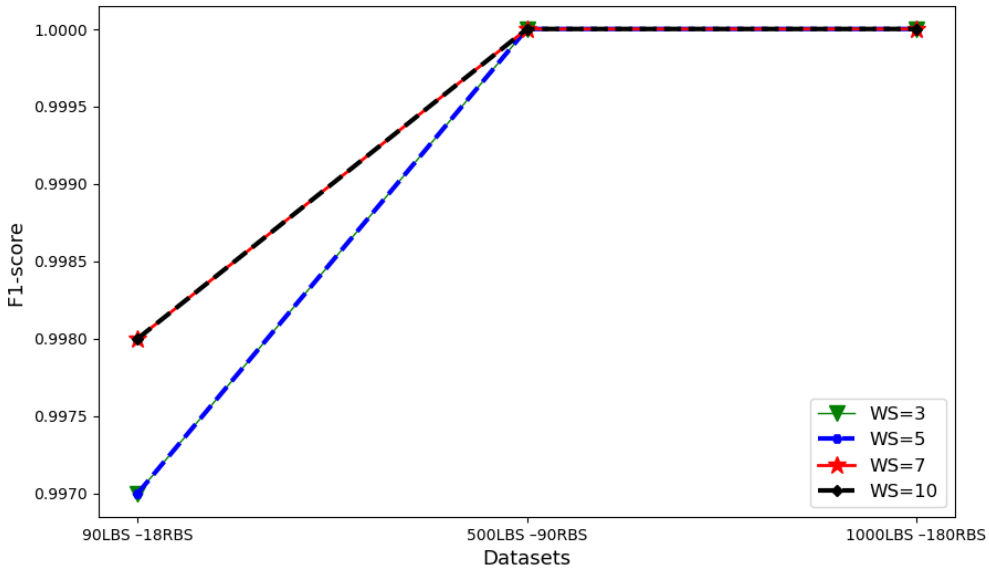


Fig. 13. F1-score metric.

589x344mm (39 x 39 DPI)

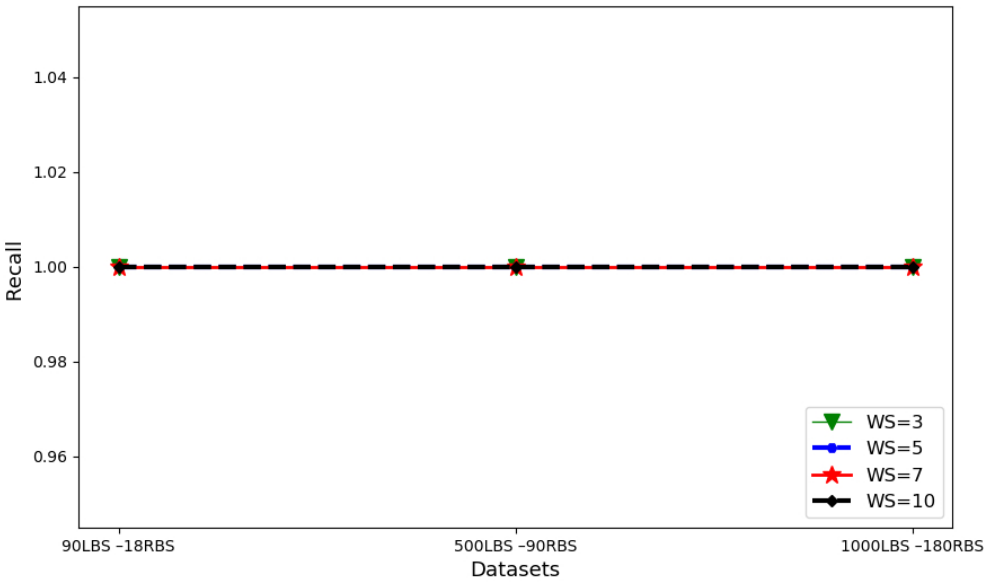


Fig. 14. Recall metric.

577x344mm (39 x 39 DPI)

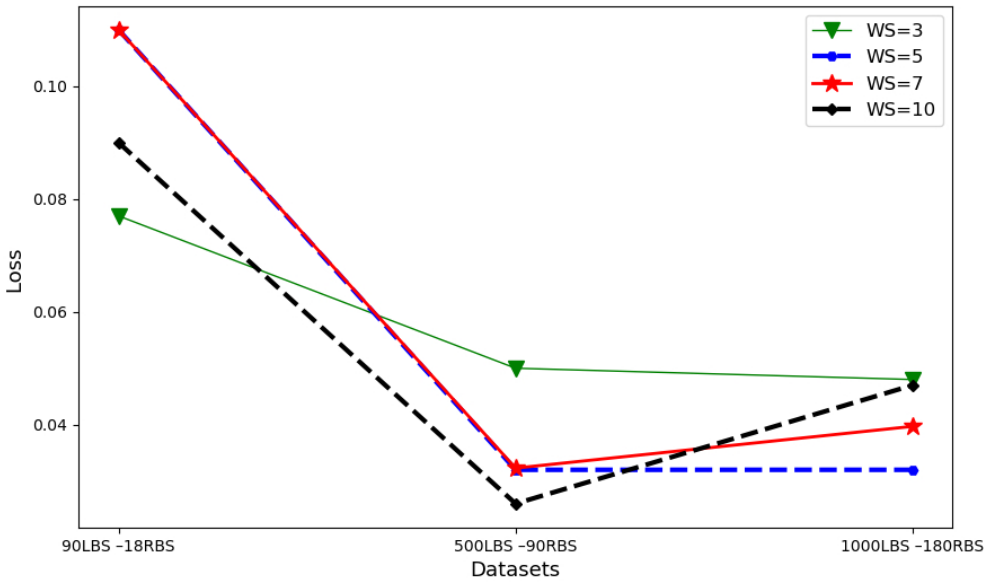


Fig. 15. Loss rate

577x344mm (39 x 39 DPI)