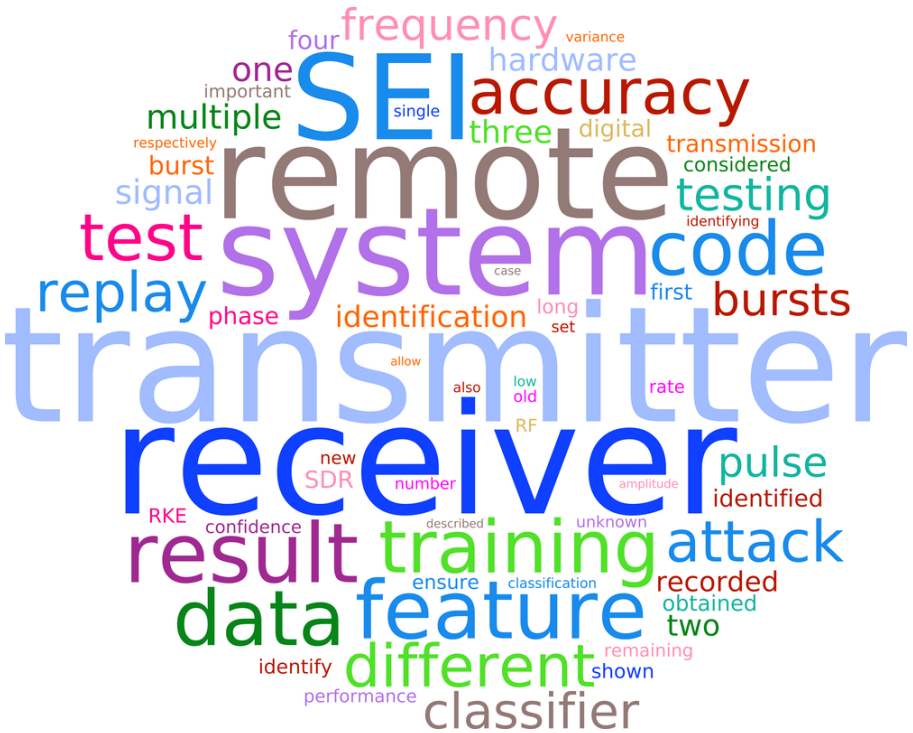
Submitted version of: Lodewicus J. Diedericks and Warren P. du Plessis, “Specific Emitter Identification with Different Transmission Codes and Multiple Receivers,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 61, no. 2, pp. 1536–1544, Apr. 2025. The published version is available online at https://ieeexplore.ieee.org/document/10669765



ABBREVIATIONS

ADC	analogue-to-digital converter
CGAN	conditional generative adversarial networks
CSIR	Council for Scientific and Industrial Research
DAC	digital-to-analogue converter
DFT	discrete Fourier transform
ELINT	electronic intelligence
ES	electronic support
EW	electronic warfare
FIR	finite impulse response
FNR	false negative rate
FPR	false positive rate
IEEE	Institute of Electrical and Electronics Engineers
ISM	industrial, scientific and medical
KNN	k^{th} nearest neighbour
MS/s	million samples per second
PCA	principal component analysis
PWM	pulse width modulation
RF	radio-frequency
RFF	radio-frequency fingerprinting
RKE	remote keyless-entry
SDR	software-defined radio
SEI	specific emitter identification
SNR	signal-to-noise ratio
SVM	support vector machine
TBD	to be determined

Specific Emitter Identification with Different Transmission Codes and Multiple Receivers

LODEWICUS J. DIEDERICKS,
University of Pretoria, Pretoria, South Africa
Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa
WARREN P. DU PLESSIS, Senior Member, IEEE
University of Pretoria, Pretoria, South Africa

***Abstract*—A specific emitter identification (SEI) system that expands previously-published results by identifying remote keyless-entry (RKE) remotes with an accuracy of over 95% even when different digital transmission codes are used is described. This system successfully rejects replay attacks with no replay attacks being incorrectly identified as known remotes. The effect of using multiple receivers is then evaluated using this SEI system. It was found that poor accuracy of under 33% was obtained when attempting to identify transmitters using an SEI system trained on data recorded by other receivers. However, including recordings from all receivers among the receivers used to provide the training data was found to increase the accuracy to over 91%. Increasing the number of receivers used to record the training data was found to slightly reduce the identification accuracy.**

***Index Terms*—Specific emitter identification (SEI), radio-frequency fingerprinting (RFF), remote keyless-entry (RKE), software-defined radio (SDR), electronic warfare (EW).**

Manuscript received 14 Feb. 2024; revised 8 Jun. 2024 and 31 Aug. 2024; accepted 4 Sep. 2024. Date of publication to be determined (TBD); date of current version 14 April 2025.

DOI: ??.????/TAES.????.???????

Refereeing of this contribution was handled by A. Person.

Lodewicus J. Diedericks is with the University of Pretoria, Pretoria, 0002, South Africa and the Council for Scientific and Industrial Research (CSIR), Pretoria, 0001, South Africa (e-mail: u16076177@tuks.co.za). Warren P. du Plessis is with the University of Pretoria, Pretoria, 0002, South Africa (e-mail: wduplessis@ieee.org).
(Corresponding author: Lodewicus J. Diedericks.)

I. Introduction

In electronic support (ES) systems, specific emitter identification (SEI), also known as radio-frequency fingerprinting (RFF), leverages the analogue characteristics of radio-frequency (RF) signals to identify individual transmitters [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. The small, hardware-induced differences between nominally-identical transmitters are too small to negatively affect the primary function of a device but can be exploited to uniquely identify each transmitter.

Development and research into SEI are important as RF systems are widely used in electronic warfare (EW), policing, and security applications [5], [6]. SEI could also be used to address security vulnerabilities in remote keyless-entry (RKE) systems [4], where reproducing digital codes through cloning and replay attacks is possible, even with rolling-code transmitters that vary their digital codes [11], [12].

Recent SEI research has largely focused on the extraction of relevant features and the classifiers that process these features to identify transmitters. For instance, semi-supervised methods utilising bispectrum feature extraction and conditional generative adversarial networks (CGAN) have been explored to improve SEI in multiple communication scenarios [8]. Additionally, time-domain transient signals have been used for radar model identification, and Bézier curve modeling of the instantaneous frequency law has been applied in the classification of radar pulses in naval contexts [9], [10].

Despite these advancements, some important issues related to SEI systems do not appear to have been considered. These include simple approaches to addressing the use of different digital transmission codes and ensuring robustness across different receiver configurations.

A limitation of some SEI systems is that they consider a single transmitter using different digital codes as multiple different transmitters (e.g. [4]). This is undesirable because changing the identification of a transmitter when a different code is used could prevent identification of that transmitter. An example of how this situation can arise is that one approach to SEI divides a burst within a transmission into equally-sized regions and then computes features for each of these regions [4], [5]. The structure of a digital code can affect the features in each region in this and similar approaches to the point that a transmitter may be misidentified.

Another significant consideration is that the effect of transferring features between different receivers in SEI

systems does not appear in the open literature. This is an important issue because there are many applications where multiple receivers have to be capable of identifying the same transmitters. For example, self-protection systems on military platforms may be called upon to identify transmitters based on information recorded by an electronic intelligence (ELINT) system, and access control requires that authorised transmitters be identified at multiple entry points.

To address the first of these limitations, an SEI system capable of identifying RKE remotes even when different digital codes are transmitted is proposed. This SEI system achieves identification accuracies of over 95% when tested with sixteen transmitters. Despite its simplicity, the SEI system does not incorrectly identify one transmitter as another in any of the tests performed. This is important in security applications where access must never be granted to the wrong transmitter [4].

The sixteen transmitters used to test the SEI system comprise ten RKE remotes purchased at the same time to ensure that they are the same make and model, three older RKE remotes that use the same codes, and three replay attacks using software-defined radio (SDR) hardware. The inclusion of replay attacks [4] that mimic two of the remotes is important because such attacks are well-known in both the security (e.g. [13]) and military (e.g. [14]) domains. This diverse set of transmitters ensures that the SEI system is comprehensively tested.

The second limitation noted above is addressed by using the developed SEI system to evaluate the effect of using different combinations of four receivers. The first tests use three of the receivers to train the SEI system and the fourth receiver to test the SEI system. The achieved identification accuracy is under 33%, which is clearly unsuitable for SEI applications. The second set of tests then consider the use of multiple receivers to both train and test the SEI system, with classification accuracies of over 91% being obtained. The test results indicate that the features of the four receivers are sufficiently similar to enable high-accuracy identification of transmitters, provided that data from all the receivers is used to train the SEI system. However, the SEI model that does not include the dataset from all receivers for training will result in poor identification accuracy.

II. Implementation of the SEI System

The implementation of the SEI system used to perform the experiments is described in this section. The overall system is similar to previously-described systems [4], [5], [7], but with modifications to allow the use of replay attacks using different hardware, different transmitter codes, and multiple receivers. The wide range of transmitters, codes, replay attack hardware, and receivers used creates a comprehensive dataset for evaluating the robustness of the developed SEI system.

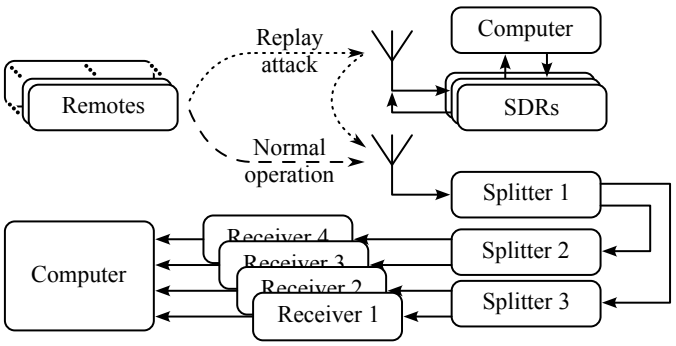


Fig. 1. The configuration of the experiments.

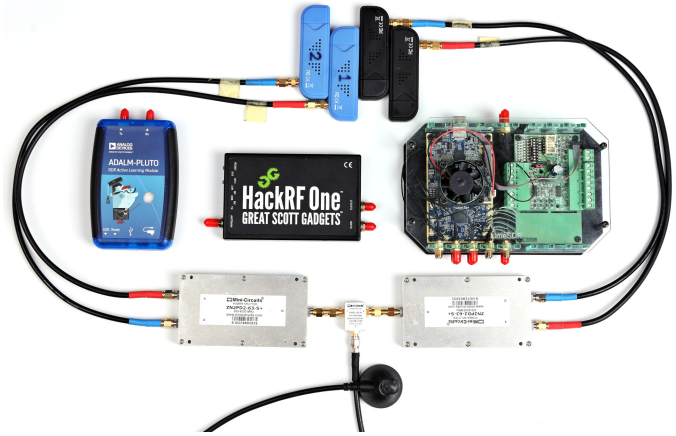


Fig. 2. The hardware used for the experiments showing the SEI receiver antenna at the bottom, the three splitters above it, the three replay-attack SDR devices between the splitters in the middle, and the four RTL-SDR receivers used by the SEI system at the top.

A. RF Hardware

Each aspect of the RF hardware will be described in this section. The configuration of the RF hardware is shown in Fig. 1, with Fig. 2 providing a photograph of the hardware used.

1) Transmitters: The transmitters that were to be identified in this study were the low-cost access remotes used in RKE systems shown in Fig. 3. The low cost of these transmitters (under $10) means that a number of identical transmitters could be purchased to test the system.

Ten gate access remotes were purchased from the same manufacturer for this study. An additional three older remotes that had been in use for some years were also included to introduce diversity. The newly-purchased remotes are designated transmitters 1 to 10,

Fig. 3. The RKE remotes used as transmitters with a new remote on the left and an old remote on the right.

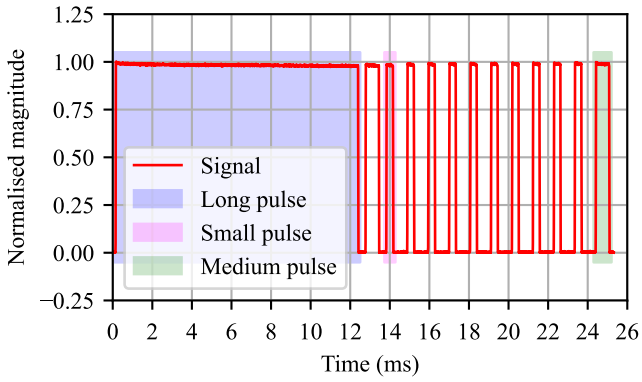


Fig. 4. Processed burst from an access remote showing the code.

while the older remotes were designated transmitters 11 to 13.

The new remotes transmit at a frequency of 434 MHz, while the older remotes transmit at 403 MHz within the relevant industrial, scientific and medical (ISM) band [4], [12]. This centre frequency difference would be easy to recreate using a simple replay attack, so this frequency difference was removed when performing SEI computations.

These remotes utilise static codes using pulse width modulation (PWM) with distinct pulse lengths to represent digital codes. For example, Fig. 4 shows a burst from remote 1 that exhibits a long pulse for transmission initiation, followed by medium and small pulses representing the code 100000000001. These bursts are then concatenated to form longer transmissions as shown in Fig. 5, with the new remotes transmitting 65 bursts each time a button is pressed, while the old remotes transmitted 85 bursts on each press of button. Two codes, 100000000001 and 101010101010, were recorded for each transmitter to test the SEI system's ability to detect the same transmitter transmitting different codes.

The remotes were measured at a distance of 3 m from the receiver antenna to prevent saturation of SEI receivers. Slight variations in this distance did occur due to practical considerations such the way the remotes were held.

2) Replay Attacks: Replay attacks were executed using GNU Radio to allow SDR hardware to record

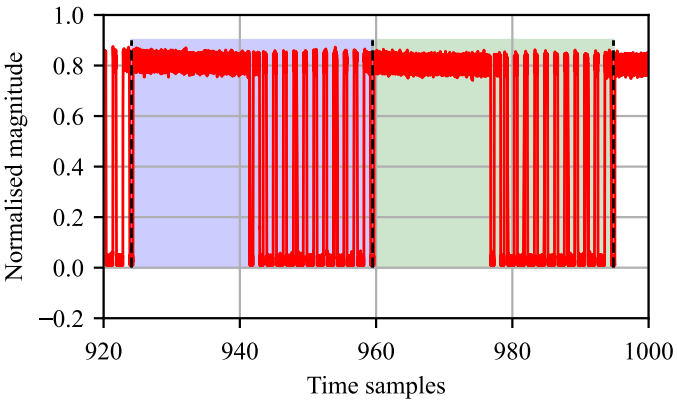


Fig. 5. Unprocessed recording of a remote indicating how bursts are combined and showing the burst boundaries.

the signals transmitted by the remotes and to replay these recorded signals using the same SDR hardware. An attempt was made to maximise the likelihood of successfully reproducing a signal by considering a range of different possibilities.

The HackRF One [15], ADALM-PLUTO [16], and LimeSDR [17] were employed for the replay attacks and are shown in Fig. 2. The HackRF One has been used before [4] but has only 8-bit analogue-to-digital converters (ADCs) and digital-to-analogue converters (DACs), while the other two devices have 12-bit ADCs and DACs, which should allow them to more accurately reproduce the signals. Each of the three SDR devices recorded both a new remote (remote 1) and an old remote (remote 11), with both remotes being recorded using both of the codes considered. Replay attacks using the HackRF One, ADALM-PLUTO, and LimeSDR are designated transmitters 14 to 16, respectively.

The signals were recorded and retransmitted using the same device parameters (e.g. centre frequency, sampling rate, gain, etc.) to minimise any differences caused by device configuration. A sampling rate of 3 million samples per second (MS/s) was used for the replay attacks because this is well above the Nyquist rate of the RKE signals while being low enough to avoid issues with missed samples and to minimise storage requirements. The gains of the SDR hardware were adjusted to ensure consistent signal-to-noise ratio (SNR) recording levels for both these SDR devices and the SEI receivers.

3) SEI Hardware: The chosen RF receiver is the RTL-SDR with an R820T tuner because the low cost of this device (around $30 [18]) means that purchasing multiple identical receivers is not prohibitively expensive. As shown in Figs 1 and 2, four such receivers were then connected to the same antenna and used to simultaneously record signals to minimise the differences

between measurements.

The sampling rate of the RTL-SDR receivers was set to 2 MS/s, which is below the maximum reliable sampling rate (2.56 MS/s [18]), but still well above the Nyquist rate of the remotes. Unfortunately, this limited bandwidth means that it is not possible to simultaneously sample both the 434 MHz and 403 MHz bands of the new and old remotes, respectively, so the centre frequency was adjusted depending on the type of remote being sampled. As noted above, this frequency difference was ignored for SEI analysis as large frequency differences are easily reproduced by replay attacks. The gain of the RTL-SDR receivers was fixed at 32 dB to ensure that the received signal did not saturate the receivers.

The four receivers were connected to a single antenna using three Mini-Circuits RF splitters which comprise one ZFRSC-123-S+ splitter (Splitter 1 in Fig. 1) and two 2N2PD2-63-S+ splitters (Splitters 2 and 3 in Fig. 1).

The SEI recordings made were 4 s long to ensure that the transmission from each remote was captured in its entirety. Each of the two codes for each remote wax captured ten times to ensure a sufficiently large number of bursts for statistical classifiers to be applied. Each remote is thus recorded a total of 20 times (ten recordings of each of two digital codes) by each receiver, giving a total of 1 300 bursts per remote for each receiver when there are 65 bursts per transmission.

B. Pre-processing

The initial steps in the signal processing involve detecting and separating individual bursts within the received signal. Subsequently, a critical aspect is removing the frequency offset to ensure accurate analysis.

The old remotes transmit more bursts per transmission than the new remotes (85 and 65 bursts, respectively), so only the first 65 bursts were used for the old remotes. This was done to ensure that there is the same amount of data for all transmitters to avoid any potential bias in the data.

The start and stop positions of each burst within the signal were identified by using the gap before the start of the long pulse to separate bursts, with an example being shown in Fig. 5. The recordings of the bursts were compiled into a dataset.

As noted above, frequency offsets are easily emulated by replay attacks, so frequency offsets were removed from all the recordings.

A finite impulse response (FIR) filter consisting of a fifty-coefficient the Blackman window was used to filter the data in an effort to reduce noise and filter out unwanted signals that may be encountered. A FIR filter was selected due to its linear phase response, which will minimise any distortion to the phase and frequency response of the transmitters [19]. The Blackman window was selected because it achieves low side-lobes without excessively broadening the passband [20].

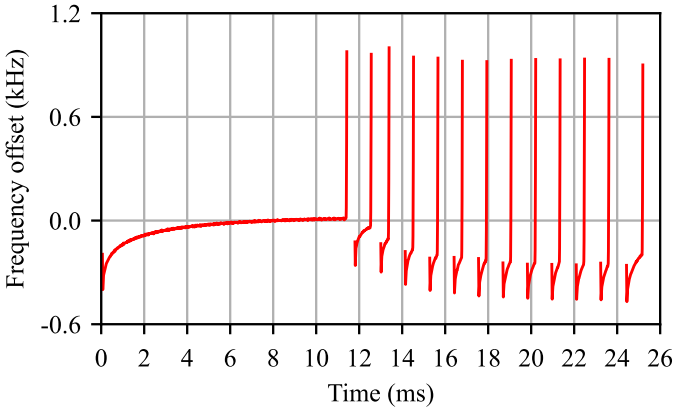


Fig. 6. The phase of a burst from remote 1 after processing.

After the filter, the amplitude of each recorded burst was normalised to ensure that there is no amplitude variance between bursts. Such amplitude variations are mainly due to changes in the environment, such as the precise positioning of the transmitter and receiver, so retaining these amplitude variations will negatively affect the performance of the SEI system.

The phase data are uniformly distributed over the range $-180°$ to $180°$ between pulses. This random phase data between pulses was discarded and only the phase of the transmissions was considered to avoid the issues this random phase would cause during feature estimation [4]. The instantaneous frequency of the received signals were used instead of the phase data to avoid difficulties with arbitrary phase offsets between pulses and phase wrapping. This was implemented by determining the gradient of the unwrapped phase. Examples of the magnitude and instantaneous frequency responses of a burst transmitted by a remote are shown in Figs 4 and 6, respectively.

C. Test and Training Data Selection

The separation of available data in to test and training sets is extremely important because it is crucial to ensure that the test data are never used during the development of a classifier as this will bias the results [21].

The first step taken to ensuring reliable testing was to use 90% of the available data for each transmitter as training data and the remaining 10% of the data for each transmitter as test data. So for example, 1 170 of the 1 300 recorded bursts for each remote would be allocated to the training data set, while the remaining 130 bursts would be allocated to the test data set.

Twice as many bursts are recorded for the replay attacks because each replay attack attempts to reproduce two different remotes, so the relevant proportions for replay attacks are 2 340 bursts and 260 bursts, respectively.

It is important to note that the testing data were only used for final testing and were not used either during training or to compare different approaches. Similarly, the comparisons between different approaches and hyperparameter optimisation were performed using the training data so as to avoid using the test data for this purpose.

The problem with using the training data for both training and testing during classifier development is that the data available are limited. K-fold cross-validation is a widely employed technique in machine learning, particularly during the development and evaluation of classifiers [21]. K-fold cross-validation serves as a valuable tool for assessing and improving classifier performance, providing a more robust and reliable estimate of a model's generalisation performance than a single train-test split [22], [23], [24]. It is important to note that the performance results obtained using K-fold cross-validation may differ from the true performance of the classifier determined using the test data, as K-fold cross-validation does not use the test data. The implementation of K-fold cross validation in the Python scikit-learn model selection library version 1.3.0 was used here, with four folds being employed.

D. Feature Extraction and Selection

Features need to be consistent for all the transmissions of a given transmitter while being distinct from the features of other transmitters [4], [5], [6]. Additionally, it would be useful if the features could be easily computed to minimise processing requirements. Finally, reducing the number of features further simplifies an SEI system by only considering those features that best distinguish the transmitters.

The use of statistical values computed for different portions of each burst as features has been shown to produce excellent results for various transmitters [3], [4], [5], including the same type of transmitters considered here [4]. The objective of this work is not to evaluate the effect of different features in SEI, so these features are regarded as acceptable for this work, especially in light of the excellent results obtained in Sections II-E and III-A.

However, the previously-used approach of considering the entire transmitted burst [3], [4], [5] results in different codes from a single transmitter being identified as different transmitters [4]. This difficulty was overcome by considering only the long pulse in each burst (see Fig. 4) to compute the features because this long pulse is common to all codes.

Previous studies divided each burst into a number of regions of equal length and computed statistical features for each of these regions [3], [4], [5]. This was found to be unnecessary here with excellent results being obtained when the entire long pulse was considered (see Sections II-E and III-A).

The statistical features used were the mean, variance, standard deviation, skewness, and kurtosis of both the magnitude and phase of the entire long pulse. An additional feature was obtained by calculating the variance of the magnitude of the discrete Fourier transform (DFT) over portions of the entire burst where the remote was transmitting (i.e. dead times between pulses were removed).

Python version 3.10 was used to calculate the statistical features and frequency variance with SciPy library version 1.11.2 functions. Notably, the signal's frequency offset is removed during the signal-processing step. The resulting dataset contains 11 features, with five for both amplitude and frequency domains, and one for the DFT variance.

The superiority of frequency data over amplitude data for classification has been demonstrated previously [4], [5]. Restricting the SEI system to the features that best distinguish transmitters will produce accurate identification while reducing the computational complexity of the system [5].

The dimensionality-reduction technique principal component analysis (PCA) was chosen for feature selection, with the ten most significant resulting features being used.

PCA is a widely used dimensionality reduction technique that transforms correlated features into uncorrelated ones, while retaining most of the original data's variance [25], [26]. Other advantages of PCA include the orthogonality of new features, noise reduction, data visualisation in lower dimensions, and highlighting influential features [27]. However, PCA has limitations such as lack of straightforward interpretation, potential information loss, and reliance on linear relationships [27], though none of these considerations are significant in this work.

E. Classification

Once the features have been calculated and managed the classifier can be implemented in an attempt to identify the individual transmitters. The objectives of this research are to allow consideration the effects of

TABLE I
CLASSIFIERS CONSIDERED FOR SEI IMPLEMENTATION

Classifier	Accuracy
Ada boost	22.2%
Decision tree	56.5%
Gradient boosting	65.2%
K^{th} nearest neighbour (KNN)	62.7%
Logistic regression	45.6%
Naive Bayes	56.9%
Neural network	50.1%
Random forest	74.2%
Support vector machine (SVM)	65.3%

different transmission codes and multiple receivers on the SEI system rather than attempting to determine the best classifier. As a result, the process outlined below is intended merely to identify a classifier that produces useful results, and is thus not a rigorous process to determine the best classifier.

The classifiers implemented in the Python scikit-learn library version 1.3.0 were considered with the default hyperparameters, and the results are shown in Table I. It can be seen that the choice of the classifier has an impact on the classification accuracy, but it is important to note that changes to the hyperparameters of each classifier could significantly affect the results.

The random forest classifier emerged as the most accurate classifiers during this initial testing, achieving an initial accuracy of 74.2%. The random forest classifier implementation used here assigns confidence scores to predictions, allowing the identification of known and unknown classes. A threshold of 60% was chosen, with predictions with confidence scores below this marked as unknown.

While promising, results obtained previously suggest that far higher classification accuracies are possible [4], so tuning of the hyperparameters of the random-forest classifier was performed. As noted above, the Python scikit-learn implementation of hyperparameter tuning using K-fold cross validation with four folds was used. For this specific dataset, using 1 000 estimators with no maximum depth limit and automatically determining the maximum features resulted in a highest accuracy of 94.5%.

III. EXPERIMENTAL RESULTS

Once the features have been calculated and managed as described above the classifier can be implemented with the resulting training and testing datasets. The first two tests in Sections III-A and III-B only consider receiver 1, while the remaining tests in Section III-C make use of all four receivers.

A. Single Transmission Code

For the first set of tests, the classifier was trained and tested with one transmission code. The primary objective of this test is validation by demonstrating that performance comparable to that obtained previously [4] is obtained.

Only 650 and 1 300 of the bursts recorded from the remotes and replay attacks, respectively, are relevant here because only one of the two codes recorded is considered.

The classifier achieved an average accuracy of 98.6%, with the lowest prediction accuracy for any of the transmitters being 95.4%. These results demonstrate the success of the SEI system in accurately identifying transmitters when trained on all 16 transmitters transmitting a single code.

The false negative rate (FNR) was 1.2%, indicating a low rate of known transmitters being identified with confidence scores too low to be reliably identified. Notably, the false positive rate (FPR) was zero, so none of the remotes was incorrectly identified as another remote.

A further test was conducted by training the classifier on the first eight transmitters, and excluding the remaining eight transmitters from the training data, with the resulting confusion matrix being shown in Table II. The SEI system demonstrated an average accuracy of 97.1% with the lowest accuracy for any of the transmitters being 95.4%. These results showcase the effectiveness of the SEI system in accurately classifying unknown remotes as such, again achieving an FPR of zero. Crucially, this outcome includes correctly identifying all replay attacks as unknown. The FNR for this scenario was 1.3%, indicating a slight increase from training using all transmitters, but still maintaining robust performance.

These results are comparable to previously-published results [4], thereby validating the SEI system.

B. Different Transmission Codes

As noted in Section II-D, the implemented SEI system uses features from only the long pulse in an effort to remove the dependency on the code used. A test was conducted using all 16 recorded remotes transmitting two different digital codes to evaluate the effectiveness of this change.

The dataset doubled in size, with the full 1 300 and 2 600 bursts for each remote and replay attack, respectively, being used here.

TABLE II
CONFUSION MATRIX WHEN USING ONE CODE AND TRAINING ON 8 TRANSMITTERS

		Predicted transmitter									
		1	2	3	4	5	6	7	8	U*	Correct
True transmitter	1	63	0	0	0	0	0	0	0	1	98.5%
	2	0	62	0	0	0	0	0	0	3	95.4%
	3	0	0	62	0	0	0	0	0	3	95.4%
	4	0	0	0	64	0	0	0	0	1	98.5%
	5	0	0	0	0	62	0	0	0	3	95.4%
	6	0	0	0	0	0	62	0	0	3	95.4%
	7	0	0	0	0	0	0	65	0	0	100%
	8	0	0	0	0	0	0	0	63	2	97.0%
	9–16†	0	0	0	0	0	0	0	0	715	100%

*Unknown (confidence score below 60%)
†Transmitters 9 to 16 were not considered during training.

TABLE III
ACCURACY RESULTS FOR MULTIPLE RECEIVERS

Training receivers				Testing receivers				Accuracy
1 (%)	2 (%)	3 (%)	4 (%)	1 (%)	2 (%)	3 (%)	4 (%)	(%)
90				10				95.5
	90				10			95.4
		90				10		95.3
			90				10	95.5
	100	100	100	100				24.1
100		100	100		100			26.3
100	100		100			100		32.8
100	100	100					100	25.6
90	90			10	10			92.8
90	90	90		10	10	10		91.9
90	90	90	90	10	10	10	10	91.2

The accuracy of the SEI system using two codes was 95.5%, with the lowest accuracy for any single remote being 83.1%. The FPR was again zero, further demonstrating that the SEI system is resistant to incorrectly identifying one transmitter as another. Although the FNR increased to 3.9% compared to the previous test, this change is largely a result of three of the 16 transmitters having accuracies of below 90%.

It is believed that this FNR increase is most probably due to the way the instantaneous frequency of the remotes appears to be affected by the length of the preceding pulses. In Fig. 6, it can be seen that the instantaneous frequency at the start of pulses reduces more slowly after a medium pulse than after a short pulse. Changing the code results in changes to the number and positions of the short and medium pulses, so it appears reasonable to assume that the instantaneous frequency of the long pulse is sufficiently affected by the code to occasionally reduce the confidence score of an identification.

Comparisons to previously-published results [4] show that the implemented SEI system is able to achieve comparable accuracy while being able to identify RKE transmitters when different digital codes are used. While there is an increase in the FNR over the published results, the FPR remains zero for the experiments performed, thereby retaining a key characteristic of the published SEI system [4].

C. Multiple Receivers

In order to test the consistency of the features across multiple receivers, various combinations of testing and training using different receivers were considered. The results for these tests are summarised in Table III.

The first set of tests was to confirm that each of the receivers produced similar results to the test described in Section III-B. The first block of results in Table III, where data from the same receiver were used for both training and testing, shows that this is indeed the case with the accuracies for the four receivers varying over the narrow range of 95.3% to 95.5%.

The next set of tests was to use data from three of the receivers to train the classifier, while using the remaining receiver to test the performance of the classifier. The results of this test are expected to be good if the features are consistent across the receivers. All the bursts from three of the receivers are used for training and all the bursts from the remaining receiver were used for testing.

The summary of the results in the second block of Table III produced extremely poor results with the accuracy varying from 24.1% to 32.8% with three of the results being 26.3% or worse. This suggests that the features are not consistent across the receivers, but further investigation is required. The confusion matrix obtained for the case with the lowest accuracy, when receivers 2, 3, and 4 were used for training and receiver 1 was used for testing, is provided in Table IV.

Unlike in the previous tests, the FPR is high in this case, which is undesirable. Additionally, there are only 58 unknown identifications, implying high confidence scores for the false-positive identifications.

TABLE IV
CONFUSION MATRIX WHEN TRAINING ON DATA FROM RECEIVERS 2, 3, AND 4, AND TESTING ON DATA FROM RECEIVER 1

		Predicted transmitter																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	U*	Correct
True transmitter	1	269	0	14	29	0	5	0	12	612	0	0	0	0	0	232	126	1	20.7%
	2	5	15	281	70	113	61	0	39	305	20	0	0	0	7	287	90	7	1.2%
	3	5	60	9	425	45	563	3	3	8	80	0	0	0	0	18	79	2	0.7%
	4	0	285	141	42	15	391	2	46	129	47	0	3	0	11	59	123	6	3.2%
	5	23	41	31	112	134	32	10	32	364	1	0	0	0	1	230	279	10	10.3%
	6	5	1	57	196	16	421	9	41	248	79	0	0	1	8	36	167	15	32.4%
	7	0	4	20	11	3	558	11	202	480	0	0	0	0	0	10	0	1	0.8%
	8	15	46	5	194	138	574	13	21	123	12	0	0	0	0	18	136	5	1.6%
	9	26	54	10	201	75	432	10	20	134	72	0	0	0	0	26	236	4	10.3%
	10	1	3	2	116	3	70	14	25	235	691	0	0	0	0	17	116	7	53.4%
	11	0	0	0	0	0	0	0	0	0	0	1250	3	0	0	47	0	0	96.2%
	12	0	0	0	0	0	0	0	0	0	0	70	1210	0	1	19	0	0	93.1%
	13	0	0	0	0	0	0	0	0	0	0	0	0	163	560	0	577	0	12.5%
	14	0	0	0	317	0	487	0	0	0	0	1003	192	28	553	20	0	0	21.3%
	15	414	0	0	69	2	96	0	4	7	2	605	0	0	0	702	699	0	27.0%
	16	349	0	37	7	0	30	0	30	330	7	849	41	0	7	39	874	0	33.6%

*Unknown (confidence score below 60%)

Perhaps more interestingly, Table IV shows that transmitters 1 to 10, the new remotes, are most often confused. These remotes are the ones that were purchased together, so their characteristics are expected to be most similar. By comparison, transmitters 11 and 12, two of the old remotes, are most often correctly identified, and transmitter 13, the remaining old remote, is not often mistaken for one of the other remotes. These observations suggest that the system is no longer performing SEI but rather transmitter classification because devices of the same make and model are no longer distinguished, while devices with different characteristics are still reasonably well classified.

In terms of replay attacks, transmitters 15 and 16, the replay attacks with the ADALM-PLUTO and the LimeSDR, are most often confused with the new remotes. Significantly, these transmitters are often identified as originating from transmitters 1 and 11, which are the transmitters used to provide data for the replay attacks. By comparison, the remaining replay attack, transmitter 14 using the HackRF One, only appears to have a significant effect on transmitter 11. This outcome is believed to be a result of the fact that the SDRs used as transmitters 15 and 16 have higher ADC and DAC resolutions of 12 bits, which suggests that their reproduction of signals should be better than the SDR used for transmitter 14, which has lower a ADC and DAC resolution of 8 bits.

Given these poor results when the receiver used for testing is not also used in training, the question now becomes whether better results are obtained if all the receivers used for testing are also used for training. The key difference in this latter case is that all receivers used for testing were also used for training, while this is not true in the former case. The results in the final block of Table III show the effect of using two, three, and all four of the receivers for both training and testing.

While the results using multiple receivers for both training and testing are lower than the results for testing and training using a single receiver in the first block of Table III, the results are all over 91%, suggesting that the system is again performing SEI.

Another observation from the final block of Table III is that the overall accuracy decreases as the number of receivers used during training increases from 95.5% for one receiver to 91.2% for four receivers. This observation suggests that the features differ between receivers, but not so much that the SEI system ceases to function when multiple receivers are used for training.

The confusion matrix when receivers 1 and 2 are used for testing and training is provided in Table V. The lowest accuracy for any of the transmitters is seen

TABLE V
CONFUSION MATRIX WHEN TRAINING AND TESTING ON DATA FROM RECEIVERS 1 AND 2

		Predicted transmitter																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	U*	Correct
True transmitter	1	248	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	95.4%
	2	0	213	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	81.9%
	3	0	0	239	0	0	0	0	0	0	0	0	0	0	0	0	0	21	91.9%
	4	0	0	0	237	0	0	0	0	0	0	0	0	0	0	0	0	23	91.2%
	5	0	0	0	0	214	0	0	0	0	0	0	0	0	0	0	0	46	82.3%
	6	0	0	0	0	0	213	0	0	0	0	0	0	0	0	0	0	47	81.9%
	7	0	0	0	0	0	0	239	0	0	0	0	0	0	0	0	0	21	91.9%
	8	0	0	0	0	0	0	0	232	0	0	0	0	0	0	0	0	28	89.2%
	9	0	0	0	0	0	0	0	0	240	0	0	0	0	0	0	0	20	92.3%
	10	0	0	0	0	0	0	0	0	0	255	0	0	0	0	0	0	5	98.1%
	11	0	0	0	0	0	0	0	0	0	0	251	0	0	0	0	0	9	96.5%
	12	0	0	0	0	0	0	0	0	0	0	0	260	0	0	0	0	0	100%
	13	0	0	0	0	0	0	0	0	0	0	0	0	260	0	0	0	0	100%
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	520	0	0	0	100%
	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	519	0	1	99.8%
	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	505	15	97.1%

*Unknown (confidence score below 60%)

to be 81.9% and the FNR is 6.0%, both of which are lower than the case where only one receiver was considered. However, the FPR is zero, so the only errors are when the confidence score is too low to reliably identify a transmitter. This desirable behaviour is thus also observed when multiple transmitters are used for training and testing.

IV. CONCLUSION

An SEI system that is capable of identifying transmitters even when different codes are used was described. This SEI system was used to investigate the effects of using multiple receivers in the SEI system.

The RF hardware of the system was configured to allow four receivers to simultaneously receive signals from thirteen RKE remotes, of which ten were purchased together. Three additional transmitters were generated by using SDR hardware to generate replay attacks mimicking two of the remotes.

The first results demonstrated that the SEI system was functioning correctly with both one and two digital codes. These results were comparable to previously-published results, but with the improvement that RKE transmitters were identified even when transmitting different codes.

When training with data from three receivers and testing using the remaining receiver, the results were poor with identification accuracies of under 33%. Further testing that used multiple receivers for both testing and training led to significantly better results with accuracies of over 91% being achieved. Finally, it was noted that the accuracy decreased from over 95% to 91% as the number of receivers used for both training and testing was increased from one to four. These outcomes suggest that the features are not identical for the different receivers, but are sufficiently similar to ensure that training with data from all receivers will lead to good results.

REFERENCES

[1] D. L. Adamy, *EW 102: A Second Course in Electronic Warfare*. Norwood, MA, USA: Artech House, 2004.

[2] A. de Martino, *Introduction to Modern EW Systems*. Norwood, MA, USA: Artech House Publishers, 2012.

[3] J. N. Samuel, "Specific emitter identification for GSM cellular telephones," Master's thesis, University of Pretoria, Pretoria, South Africa, Jun. 2017.

[4] J. N. Samuel and W. P. du Plessis, "Specific emitter identification for enhanced access control security," *SAIEE Afr. Res. J.*, vol. 108, no. 2, pp. 71–79, Jun. 2017.

[5] D. Reising, T. M. A, and M. M. J., "Improved wireless security for GMSK-based devices using RF fingerprinting," *Int. J. Electron. Secur. Digit. Forensics*, vol. 3, no. 1, pp. 41–59, Mar. 2010.

[6] K. I. Talbot, P. R. Duley, and M. H. Hyatt, "Specific emitter identification and verification," *Technol. Rev. J.*, vol. 113, pp. 113–133, Jan. 2003.

[7] S. Deng, Z. Huang, and X. Wang, "A novel specific emitter identification method based on radio frequency fingerprints," in *IEEE Int. Conf. Comput. Intell. Appl. (ICCIA)*, Beijing, China, 8–11 Sep. 2017, pp. 368–371.
[8] K. Tan, W. Yan, L. Zhang, Q. Ling, and C. Xu, "Semi-supervised specific emitter identification based on bispectrum feature extraction CGAN in multiple communication scenarios," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 1, pp. 292–310, Feb. 2023.
[9] S. Guo, S. Akhtar, and A. Mella, "A method for radar model identification using time-domain transient signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 5, pp. 3132–3149, Oct. 2021.
[10] F. Digne, A. Baussard, A. Khenchaf, C. Cornu, and D. Jahan, "Classification of radar pulses in a naval warfare context using Bézier curve modeling of the instantaneous frequency law," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1469–1480, Jun. 2017.
[11] K. Greene, D. Rodgers, H. Dykhuizen, K. McNeil, Q. Niyaz, and K. A. Shamaileh, "Timestamp-based defense mechanism against replay attack in remote keyless entry systems," in *IEEE Int. Conf. Consum. Electron. (ICCE)*, Las Vegas, NV, USA, 4–6 Jan. 2020, pp. 1–4.
[12] S. van de Beek and F. Leferink, "Vulnerability of remote keyless-entry systems against pulsed electromagnetic interference and possible improvements," *IEEE Trans. Electromagn. Compat.*, vol. 58, no. 4, pp. 1259–1265, Aug. 2016.
[13] P. Syverson, "A taxonomy of replay attacks [cryptographic protocols]," in *Computer Security Foundations Workshop*, Franconia, NH, USA, 14–16 Jun. 1994, pp. 187–191.
[14] P. E. Pace, *Developing digital RF memories and transceiver technologies for electromagnetic warfare*. Norwood, USA: Artech House, 2022.
[15] (2024, 30 May) HackRF One – Great Scott Gadgets. [Online]. Available: https://greatscottgadgets.com/hackrf/one/
[16] (2024, 30 May) ADALM-PLUTO evaluation board – Analog Devices. [Online]. Available: https://www.analog.com/en/resources/evaluation-hardware-and-software/evaluation-boards-kits/adalm-pluto.html
[17] (2024, 30 May) LimeSDR – Lime Microsystems. [Online]. Available: https://limemicro.com/products/boards/limesdr/
[18] (2024, 30 May) About RTL-SDR. [Online]. Available: https://www.rtl-sdr.com/about-rtl-sdr/
[19] M. A. Parker, *Digital Signal Processing 101: Everything You Need to Know to Get Started*. Boston, NY, USA: Elsevier Science & Technology, 2010.
[20] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.
[21] S. J. Russell and P. Norvig, *Artificial intelligence: A modern approach*, 4th ed. Harlow, UK: Pearson, 2022.
[22] T.-T. Wong and P.-Y. Yeh, "Reliable accuracy estimates from k-fold cross validation," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 8, pp. 1586–1594, Apr. 2020.
[23] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 569–575, Mar. 2010.
[24] S. Yadav and S. Shukla, "Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification," in *IEEE Int. Conf. Adv. Comput.*, Bhimavaram, India, 27–28 Feb. 2016, pp. 78–83.
[25] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *Sci. Inf. Conf.*, London, UK, 27–29 Aug. 2014, pp. 372–378.
[26] P. Uddin, A. Mamun, and A. Hossain, "PCA-based feature reduction for hyperspectral remote sensing image classification," *IETE Tech. Rev.*, vol. 38, no. 4, pp. 377–396, 2021.
[27] P. A. Kumari and G. J. Suma, "An experimental study of feature reduction using PCA in multi-biometric systems based on feature level fusion," in *Int. Conf. Adv. Electr., Electron. Syst Eng. (ICAEES)*, vol. 16, 14-16 Nov. 2016, pp. 109–114.

Lodewicus J. Diedericks completed the B.Eng. (Electronic), and B.Eng.Hons (Electronic) degrees from the University of Pretoria in 2020, and 2021 respectively. He is currently an employee at the Council for Scientific and Industrial Research (CSIR) and his primary research interests are related to specific emitter identification (SEI) and electronic warfare (EW).

Warren P. du Plessis (M'00, SM'10) received the B.Eng. (Electronic), M.Eng. (Electronic), and Ph.D. (Engineering) degrees from the University of Pretoria in 1998, 2003, and 2010, respectively, winning numerous academic awards including the prestigious Vice-Chancellor and Principal's Medal. He is an Associate Editor of the IEEE Transactions on Aerospace and Electronic Systems.

He spent two years as a lecturer at the University of Pretoria, and then joined Grintek Antennas as a design engineer for almost four years, followed by six years at the Council for Scientific and Industrial Research (CSIR). He has been with the University of Pretoria since December 2012, where he holds the rank of Professor. His primary research field is EW with an emphasis on cross-eye jamming, but he also conducts research on other topics including single-sensor imaging and engineering education.