A Machine-Learning Based Wireless Intrusion Detection System

Jeffrey L. Duffany. Ph.D., Carlos Y. Velez
Universidad del Turabo , Politechnic University Informatics,
jeduffany@suagm.edu, velez.carlos.y@gmail.com

Abstract

One method of designing an Intrusion Detection System (IDS) is to have it first learn what is normal behaviour and then program it to send an alert whenever a significant deviation occurs. A Software Defined Radio (SDR) system can scan a desired range of the radio frequency spectrum and obtain periodic snapshots of communication activity for example WiFi spectra. Machine learning techniques can then be applied to detect anomalous activity over the airwaves. Cluster analysis demonstrated the possibility of developing an IDS using a wavelet transform to convert WiFi spectra into a feature vector comprised of the energy in each wavelet component. Training sets of 100 WiFi spectra were used, some with intrusions and others not. The R Language was used to implement the prototype IDS using both the Naive Bayes classifier and Neural Networks.

Introduction

Intrusion Detection Systems (IDS) employ a variety of techniques to detect unauthorized use of network resources (Scarfone 2010). These automated systems can be ever vigilant and can sometimes see things that are not obvious or evident to the human eye. They can even self-adapt to an ever evolving normal behaviour and are capable of detecting transient threats that try to slip under the radar by appearing briefly and at random. For example video surveillance of a building requires a device to capture the signal and some method of pattern recognition. Historically this has been done by human beings such as a security guard however it could also be performed by some kind of automated system. It might be desirable in some cases to deploy more than one such system to add layers of security that opponents may not suspect the existence of.

Wireless signals are invisible to the human eye therefore some type of hardware and software is required to capture and display them. There are many devices currently in existence for doing this unfortunately most of them are specific to a particular wireless system. In the last few years very inexpensive Software Defined Radio (SDR) systems (Dillinger 2003) have been developed than can capture signals across a wide range of the radio frequency spectrum from DC to upwards of 3GHz. These devices can serve as an inexpensive front end to deliver the raw signals required by an intrusion detection system.

Machine learning (Lantz 2015) works by first classifying a large number of known signals and using that information to classify unknown input signals. For this to be possible it needs an input known as a feature vector which measures one or more characteristics of the raw input signal. Recently there has been research about using wavelet transforms (Akansu 2009) and machine learning on electrocardiogram and electroencephalogram signals (Amin 2015, Karpagachelvi 2011). For those applications machine learning was able to detect the difference between a normal and abnormal signal. The idea for the current research was to apply this idea to SDR radio frequency signals. The idea is detect any abnormality or intrusion as if it were some kind of off-normal or medical condition.

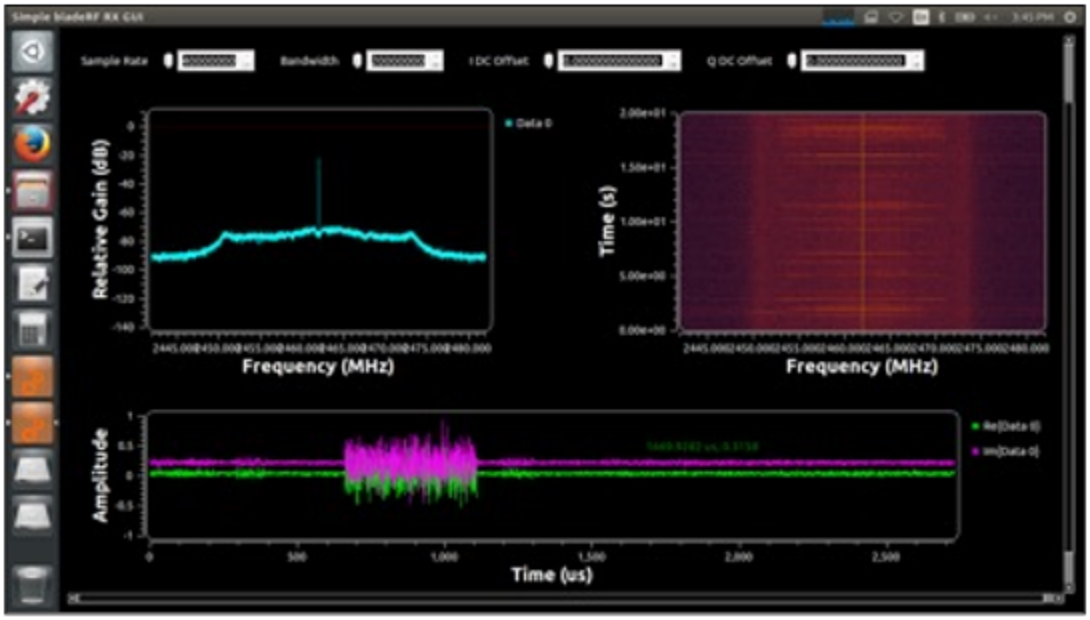


Figure 1: GNU Radio Companion Graphical User Interface

SDR Hardware and Software Platforms

Software defined radio systems are very inexpensive and versatile and are often no more than the size of a USB thumb drive. Despite their small physical size they are capable of acquiring signals from across the radio frequency spectrum. Free software is widely available to convert the captured RF signals into various analog and digital forms for example the GNU radio companion shown in Figure 1. Many of the least expensive models can only receive signals while the more expensive models can also transmit. SDR can be used to listen to AM or FM radio, obtain GPS information from commercial airliners or even control a ceiling fan. Figure 1 shows a 22 MHz wide frequency spectra window centered around WIFI channel 11 in the upper left-hand corner. In the upper right of the figure we see the waterfall diagram which shows how the WiFi spectrum varies over time. At the of the display bottom is seen an individual WiFi packet in the time domain. This screen is configurable by the user.

SDR-Based Intrusion Detection System

Based on observation of the WiFi frequency spectra using the SDR hardware we decided to build a prototype intrusion detection system using machine learning. This was done as a proof of concept to determine if the wavelet transform approach was feasible. The R Language was chosen to implement the prototype IDS because of the powerful machine learning capabilities that are natively built into the R Language platform (Lantz 2015). The baseline case is to have information actively flowing on channels 1, 6 and 11. An intrusion would be represented by significant energy on any of the other channels. The idea is to see if the difference would show up in the wavelet transform. If it does then there is a potential that this could be used as a basis for an intrusion detection system.

Wavelet Transforms

A wavelet is a small wave of finite duration that can be used in signal processing to decompose an arbitrary signal into a set of components similar to a Fourier Transform (Akansu 2009). These components contain the information of the original signal and can be used to reconstruct the original signal. There are a number of different types of wavelet transform each corresponding to a different wavelet function. The most well known are perhaps the Haar and the Daubechies wavelets (Akansu 2009). Figure 2 shows the decomposition of a typical WiFi spectra into three subwavelets (W1,W2,W3). There are limits to the number of subwavelets that can be computed for a signal depending on the number of signal samples which should ordinarily be a power of two (we are using 1024 samples).

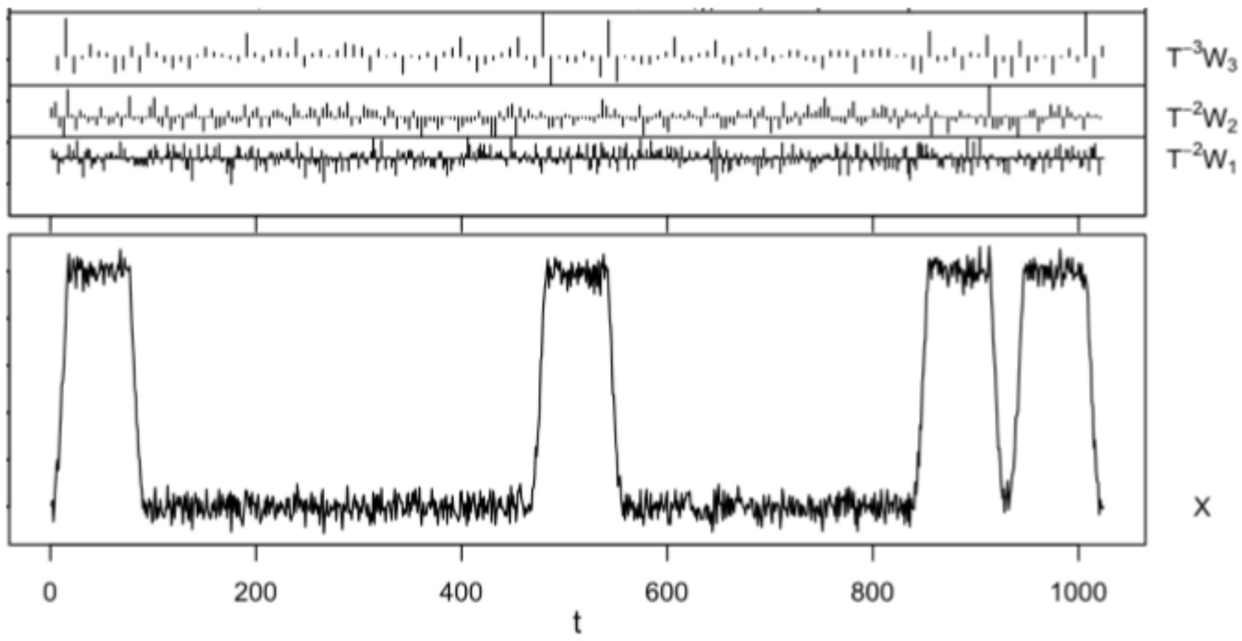


Figure 2: WIFI Spectral Density and Wavelet Components W1, W2 and W3

The lower part of Figure 2 shows a WiFi spectrum with activity on channels 1, 6, 10 and 11 and the upper part of the figure shows the wavelet components computed by R Language (Lantz 2015). The energy in each wavelet component is used as input to the machine learning algorithms. The WiFi spectrum in Figure 2 represents an intrusion as we have previously stated our baseline case only assumed communications on channels 1,6, and 11.

Machine Learning

Machine Learning attempts to create a model for a set of training data that can be used to classify or make predictions regarding unknown future data (Lantz 2015). Machine learning includes techniques such as Naive Bayesian and Neural Networks. Naive Bayesian analysis uses the joint probabilities of a set of variables to estimate whether a particular set of conditions most likely exists given the observations (Lantz 2015). The Naive Bayesian is called naive because it assumes independence of the variables and is simple to compute. It often gives good results despite the assumption of variable independence not being true. The variables act as predictors of an underlying state or condition. Neural networks use a simplified model of the neurons in the human brain and a set of training data to learn how to recognize patterns (Lantz 2015). Neural networks can be rather computationally intensive but can perform better than Naive Bayesian analysis.

Cluster Analysis

Our methodology was to employ several data mining and machine learning techniques mainly cluster analysis, Naive Bayes analysis and Neural Networks. For the cluster analysis the R Language was used on 100 WiFi spectra some of which had intrusions and others that did not. We tried several different wavelet transforms including the Haar, Daubechies and Coiflet since they are built into the R Language. We also tried decomposing the signal into different numbers of wavelet components starting with 3 (W1,W2,W3) and increasing to 7. The energy (Xi) in each of the sub-wavelets (Wi) was computed as the sum of the squares of each element of the wavelet vector.

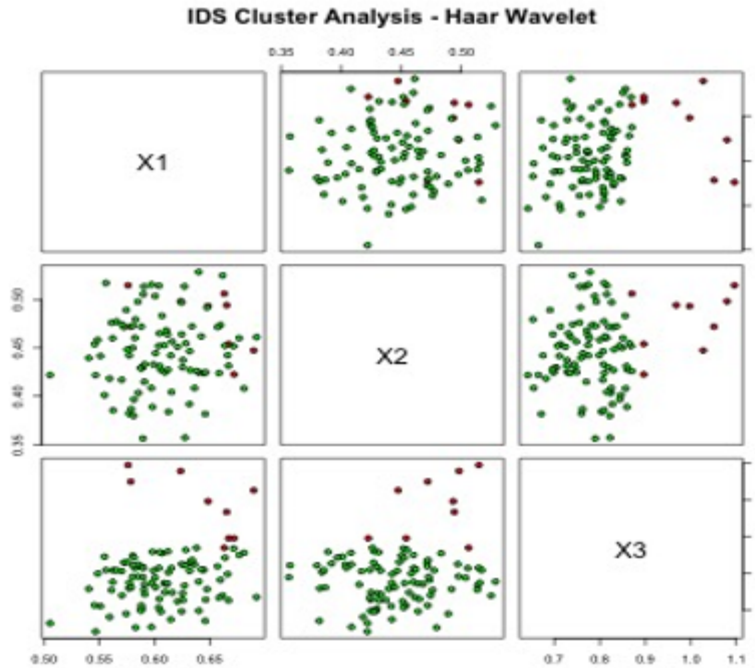


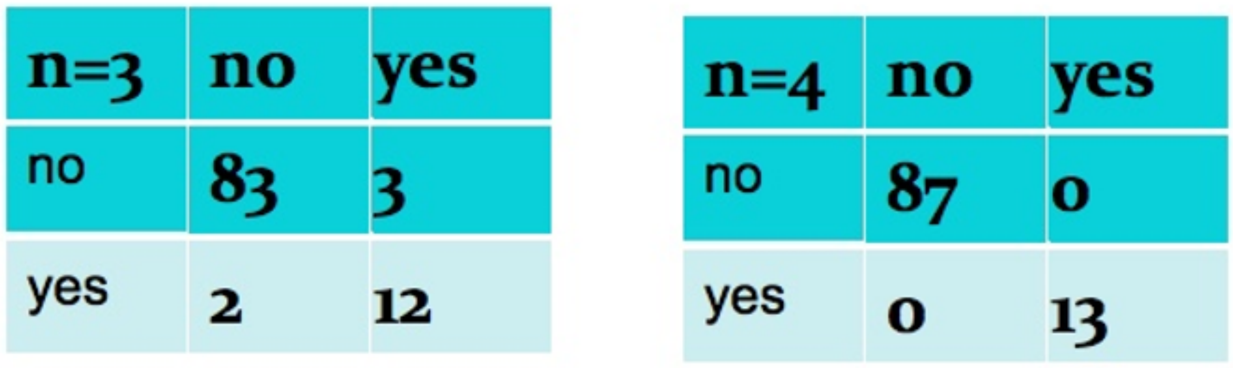
Figure 3: Cluster Analysis of energy in 3 wavelet components

Figure 3 shows the result of the cluster analysis for 100 WiFi frequency spectra using wavelets. The WiFi spectra with no intrusion are shown as green dots while the WiFi spectra with intrusions are shown in red in Figure 3. There appears to be a significant difference between the intrusion spectra and the non-intrusion spectra and particularly for the W2 and W3 wavelets. The W1 wavelet consists mostly of high frequency components and noise. The W2 and W3 wavelets contain mostly mid-band and lower frequency information, respectively. The visible clustering in Figure 3 is a indication that the wavelet transform approach has the potential to provide useful machine learning results. Use of different wavelets did not appear to improve the clustering. However use of more subwavelets did appear to improve the separation between the different spectra.

Naive Bayes Analysis

The result of the Naive Bayesian analysis is shown in Figure 4 for 100 WiFi spectra. Naive Bayesian analysis often gives a good indication about how well other machine learning techniques might work on the same data. The left column shows whether or not there was an intrusion in the WIFI spectrum (no= no intrusion, yes=intrusion). The top row shows the result of the Naive Bayes classifier (no= no intrusion, yes=intrusion).. The left table shows the result for n=3 wavelets and the right table shows the results for n=4 wavelets. The results shown are for the Haar wavelet. Use of other wavelets gave similar results.

As can be seen from Figure 4 the Naive Bayes classifier was correct in identifying an intrusion most of the time. For n=3 subwavelets it correctly predicted there was no intrusion 83 times out of the 86 times that there was no intrusion (see Figure 4 n=3 table). There were 3 false positives when the classifier said there was an intrusion when there wasn't. A small percentage of false positives is considered acceptable and perhaps even desirable in an intrusion detection system (Scarfone 2010). Of the 14 times there was an intrusion it correctly predicted the intrusion 12 times out of 14 times. There were 2 false negatives where the classifier incorrectly said there was no intrusion when there was. The design of an intrusion detection system should attempt to minimize false negatives. In summary Naive Bayes Classifier was correct 95% of the time for n=3 wavelets. When the analysis was repeated for n=4 wavelets the Naive Bayes intrusion classifier was correct 100% of the time.



n=3	no	yes
no	83	3
yes	2	12

n=4	no	yes
no	87	0
yes	0	13

Figure 4 Result of Naive Bayes Analysis for WiFi Intrusion Detection Classifier

Alternatives to Wavelet Transforms

There are alternatives to using the wavelet transform for an intrusion detection based on machine learning. One alternative is to use the Fast Fourier Transform instead of the wavelet transform. Another alternative is to use the raw unprocessed WiFi spectra. To use these approaches a decision would have to be made on how to split the frequency bands to create a feature vector. The wavelet transform approach does not suffer from this limitation.

The R Language Programming Environment

The R Language (Lantz 2015) provided the software tools used for the signal analysis. R is a high-level programming language that is especially known for its data mining and machine learning capabilities. The SDR equipment is acting like a front end to capture waveforms for analysis. Data can be output from the SDR into a CSV (comma separated value) file that can be imported into the R language environment. The "scan" command reads data from a CSV file in the working directory. This can be imported into a vector "x" in R as follows using the scan function:

```
x<-Mod(scan(file="sdr.data", what=complex, sep="\n"))
```

R uses the built-in function "dwt" to compute the discrete wavelet transform. Here the discrete wavelet transform is computed and stored in the variable "wt".

```
wt<-dwt(x, filter="haar", n.levels=3, boundary="periodic")
```

Neural Network Analysis

Neural Networks try to simulate the operation of the human brain with neurons replaced by computational units known as TLUs (Threshold Logic Units). The design of the neural network varies but normally has an input layer, output layer and one or more hidden layers. A fully connected neural network was created with 3 inputs, 2 hidden layers and one output using R Language and the result is shown in Figure 5. The output was labelled IDS and was trained to be IDS=0 for no intrusion and IDS=1 for intrusion.

Three wavelet energy components (X1,X2,X3) from 3 subwavelets (W1,W2,W3) were used to train the neural network. The neural network was trained for 100 WiFi spectra where some had intrusions and others did not. The result was that neural network was successfully able to distinguish intrusions with 100% accuracy. This can be compared with the Naive Bayesian classifier which required 4 wavelet components (W1,W2,W3,W4) to reach 100% intrusion detection accuracy.

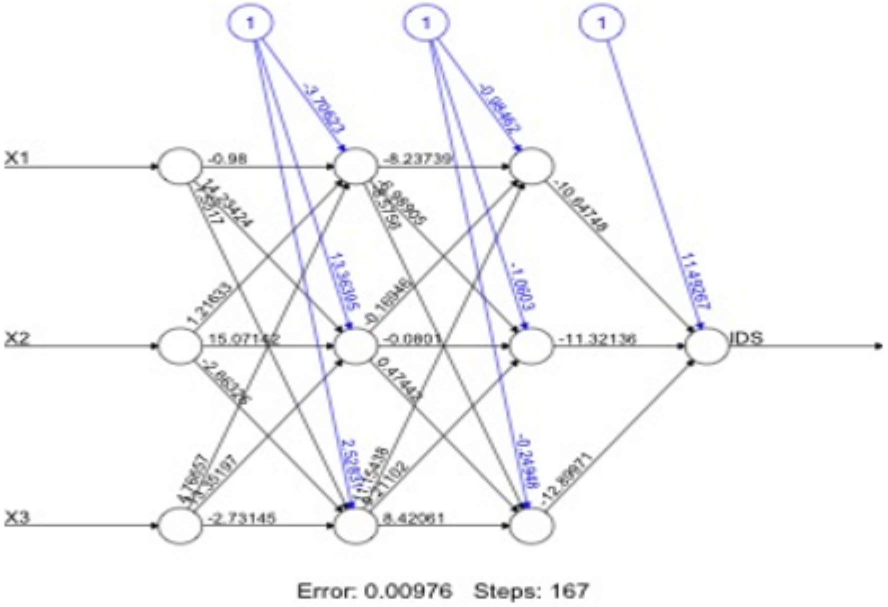


Figure 5: Neural Network Wireless Intrusion Detection System

Discussion and Future Work

One of the major cybersecurity challenges is to protect against unknown threats. For example machine learning is already being used to detect malware that has never been seen before. Results of this investigation suggest there could be some value in the wavelet approach but it needs more realistic evaluation.

We tried to simulate as realistic as a situation as possible as a proof of concept. However to develop the idea further we would need to scan a larger portion of the RF spectrum which could include the entire SDR bandwidth from DC to 3.3GHz. Also to continue to use supervised learning we would need to identify what an actual intrusion looks like Unfortunately that is a relatively rare event and therefore hard to train for. The other alternative is to use unsupervised learning. In other words train the system for a day or two assuming that there is no intrusion and then create an artificial intrusion and see if it recognizes it as abnormal. Whether this approach can be used in a practical system yet remains to be seen and would require further testing.

It is not known how accurate this approach might be in predicting intrusion. However it does not have to be perfectly accurate to be useful. False alarms are acceptable especially if valid intrusions are detected with high probability. In other words false positives are acceptable also long as there are not an excessive amount of false negatives.

A method needs to be devised to determine what advantage the wavelet approach has over other techniques. A prototype IDS should be built using wavelets and machine learning to test the concept in a practical environment across different regions of the RF spectrum.

Summary and Conclusions

A prototype wireless intrusion detection system was investigated based on the use of machine learning and RF signals obtained from Software Defined Radio to detect anomalous activity over the airwaves. The approach was based on recent research showing that wavelets (Akansu 2009) and machine learning (Lantz 2015) could be used to diagnose certain medical conditions (such as arhythmia) from analysis of time domain ecg (electrocardiogram) and eeg (electroencephalogram) signals (Amin 2015, Karpagachelvi 2011). One of the main difficulties in machine learning is reducing the dimensionality of the input. Our solution was to use the WiFi spectra from the SDR and decompose them into a small number of wavelets Wi using the wavelet transform. The energy Xi of each wavelet Wi was computed and used to create the required feature vector. Cluster analysis revealed that there was an ability to distinguish between WiFi signals in a way that might be applied to an intrusion detection system. This was corroborated by the subsequent Naive Bayes and Neural Network analyses. The R language (Lantz 2015)was used to rapidly prototype an intrusion detection system using its built-in capability for computing wavelet transforms and performing cluster, Naive Bayes and Neural Network analyses.

Acknowledgements

We would like to thank the Air Force Research Lab (AFRL) in Rome, New York for giving us the opportunity to conduct SDR research. We especially would like to thank our mentors Dr. Alfredo Vegairizarry and Dr. Tomasz Zanko for all of their advice and support throughout the investigation.

References

Akansu, A. N., Serdijn, W. A. and Selesnick, I. W. 2009. Emerging applications of wavelets: A review. Physical Communication. 3: 1. doi:10.1016/j.phycom.2009.07.001

Amin H. U., Malik A.S., Ahmad R.F., Badruddin N., Kamel N., Hussain M., Chooi W.T. 2015. Feature Extraction and Classification for EEG signals using Wavelet Transform and Machine Learning Techniques, Australas Phys Eng Sci Med. 2015 Mar;38(1):139-49. doi: 10.1007/s13246-015-0333-x.

Dillinger, M., Madani, K. and Alonistioti, N. 2003. Software Defined Radio: Architectures, Systems and Functions, Wiley & Sons, 2003, ISBN 0-470-85164-3.

Karpagachelvi, S., Arthanari, M., Sivakumar, M., 2011. Classification of Electrocardiogram Signals With Extreme Learning Machine and Relevance Vector Machine, International Journal of Computer Science Issues, Vol. 8, Issue 1, January 2011.

Lantz, B. 2015. Machine Learning with R, Second Edition, ISBN-13: 978-1784393908, Packt Publishing, July 31, 2015.

Scarfone, K. and Mell, P. 2010. Guide to Intrusion Detection and Prevention Systems (IDPS), Computer Security Resource Center, National Institute of Standards and Technology, 800-94, January 2010.