A New Intrusion Detection System Based on RQA

Dr. Bahaa Al-Musawi

bahaa.almusawi@uokufa.edu.iq

Faculty of Engineering, University of Kufa

About me

- 2018: PhD in Telecommunication, Swinburne University of technology Australia. “Detecting BGP Anomalies Using Recurrence Quantification Analysis”
- 2020-present: Doctor of Computer Network, University of Kufa
- 2020-2021: Research fellow with Mutually Agreed Norms for Routing Security (MANRS), a global initiative supported by the Internet Society
- 2018: Research assistant for project “Enterprise Network Routing Security”, Swinburne University of Technology, A project supported by DATA61 and MoD, Australia
- 2016: APNIC Internet Operations Research Grants under the ISIF Asia grants 2016 scheme, for a project titled "Rapid detection of BGP anomalies“. ~$28K
- 2015 Granted access to Virtual Internet Routing Lab (VIRL) / Cisco under academic license
- 2015 A partial PhD stipend from Cisco USA (URP grant) ~$81K

Outline

- Introduction
- Recurrence Quantification Analysis (RQA)
- The Proposed System Design
- Results and Discussion
- Conclusions

Introduction

- IoT-based systems are attacked by multiple threats such as DDoS, probing, and different information gathering attacks
- These threats are continuously increased and continually growing in sophistication
- Inspection packet's payload is impractical with high-speed networks, and it fails if the packet is encrypted
- Cyber-attacks can encrypt their communications to evade detection
- The conventional approach for building a system that can detect these threats is to inspect the packet's payload

Introduction

- The alternative approach implies studying the overall behavior of the set of packets that pass through the network (Flow-based Intrusion IDSs)
- Flow-based Intrusion IDSs:
  - Signature-based IDSs
  - ***Anomaly-based IDSs***
- Building robust anomaly-based IDSs has many challenges:
  - Impossible to capture all possible normal behaviors
  - Huge amount of training and testing data
  - Speeds of the networks today
  - The need to deal with a large number of features

Motivation

- Develop a new anomaly-based IDS with:
  - A high detection accuracy
  - Using a minimal number of features
  - Does not require a large amount of historical data
- Consequently, we introduce a new IDS based on using RQA
  - RQA is non-linear statistical analysis method based on the concept of phase plane trajectory
  - RQA shows can discover hidden patterns extracted from a single feature that may not be discovered by analyzing the one-dimensional system behavior

Recurrence Quantification Analysis (RQA)

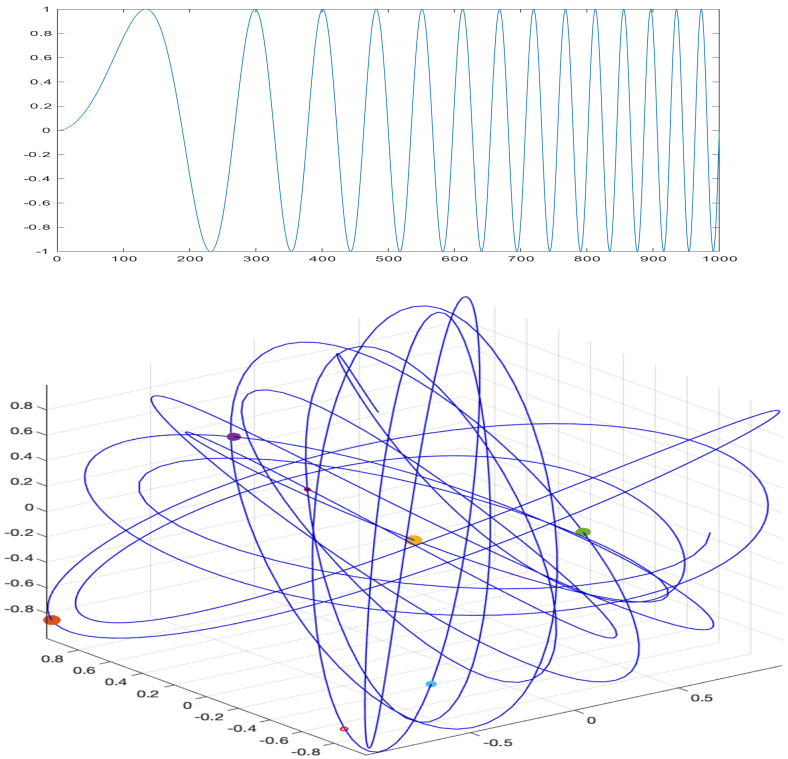

Figure (1) Phase Space trajectories for cosine function with constant amplitude and variable frequency

Recurrence Quantification Analysis (RQA)

- To enable users to investigate m-dimensional phase space trajectory using a two-dimensional representation, Recurrence Plot (PR) has been introduced
- RP is an advanced non-linear analysis technique

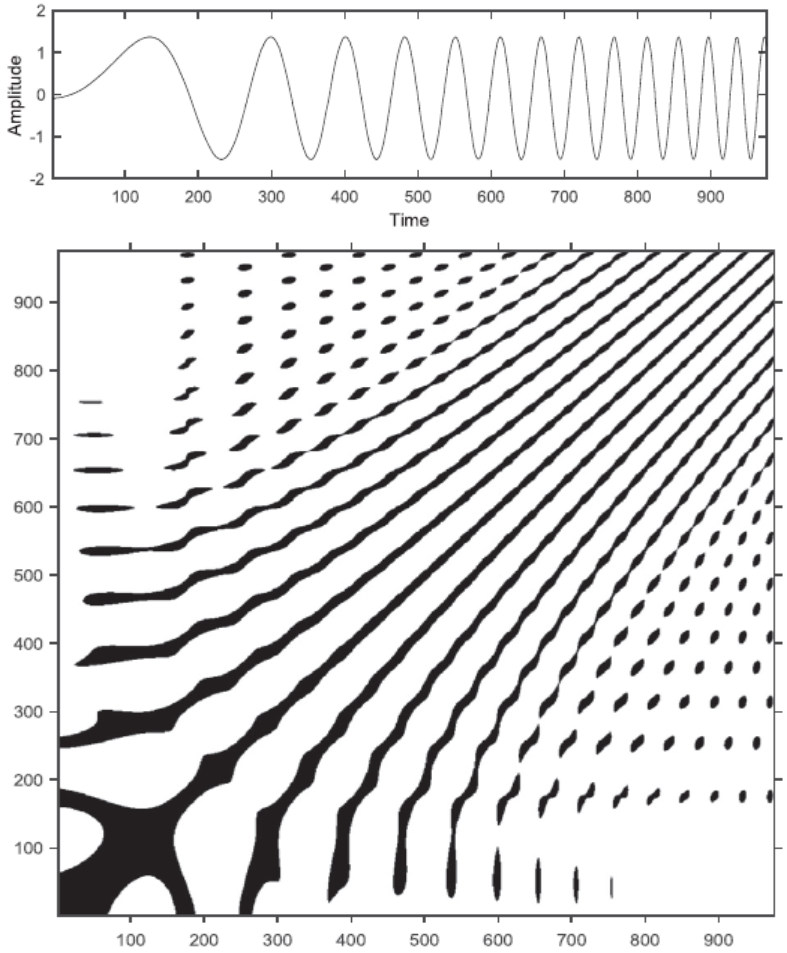

Figure (2) RP of a Cosine function with constant amplitude and variable frequency

Recurrence Quantification Analysis (RQA)

- Interpreting an RP requires a high level of experience especially for complex data
- RP cannot be directly used for real-time anomaly detection
- Consequently, RQA has been introduced to provide several measures of complexity
- RQA has multiple measurements called RQA measurements

Recurrence Quantification Analysis (RQA)

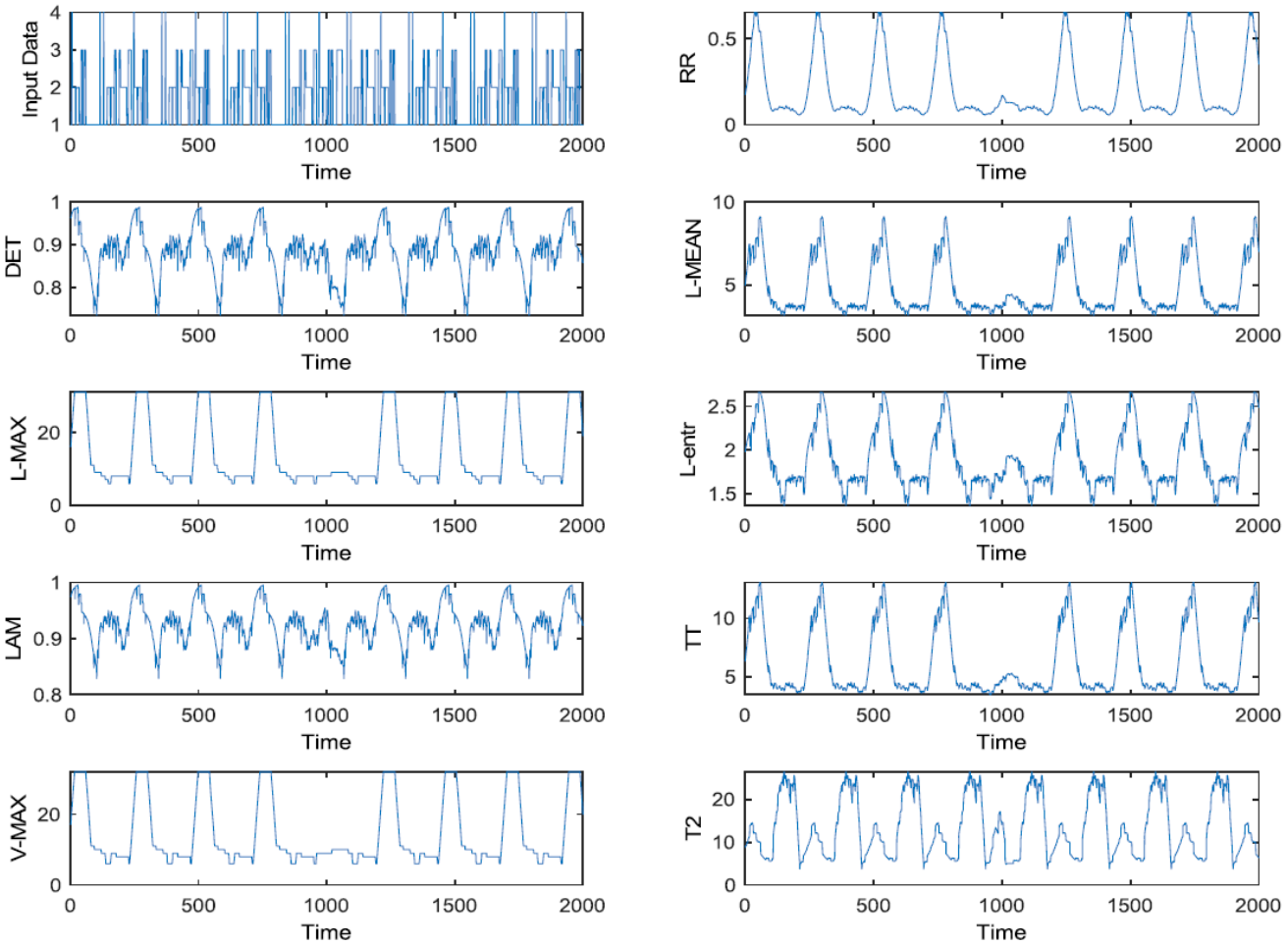

Figure (3) The effect of the sudden change in the recurrence pattern at 1040–1060 s on the RQA measurements

The Proposed System Design

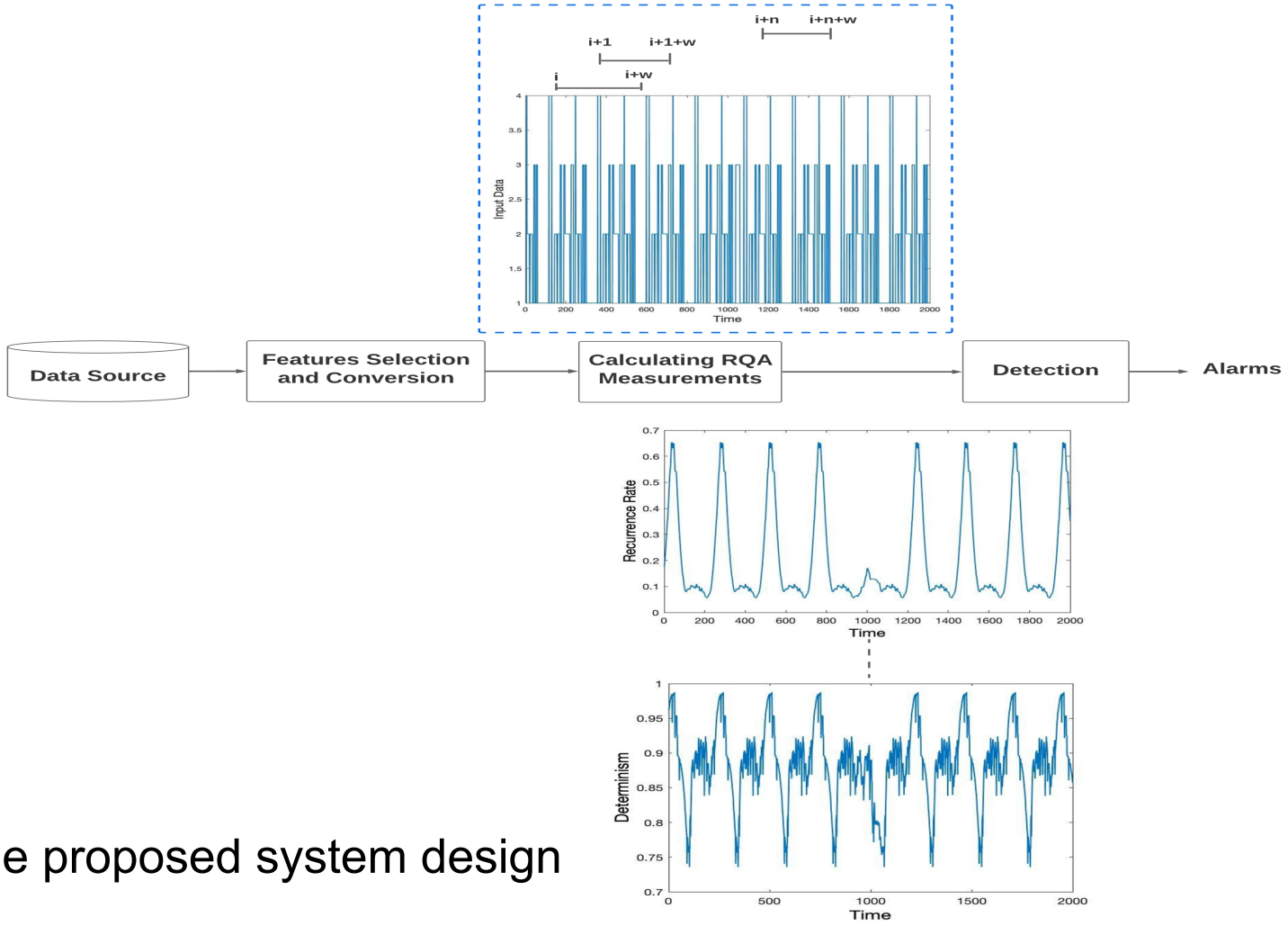

Figure (4) The proposed system design

Feature Selection Process

Table (1) The resulted features from the selection process

Feature	Description
sport	Source port number
state	The state and its dependent protocol
dsport	Destination port number
srcip	Source IP address
dstip	Destination IP address
ct_srv_src	The number of connections with the same service and the same source IP address during 100 connections
sbytes	Number of bytes sent from the source

	feature	score
1	sport	8596.724624
4	state	6645.776761
2	dsport	6156.110461
0	srcip	3242.068511
3	dstip	252.478770

Results and Discussions

Reference	Year	Method	Classifiers	Accuracy	Sensitivity	No. of Features
This work	**2022**	**RQA**	**LR, KNN, DT, and RF**	**96.71%**	**0.972**	**1**
Our Previous work [8]	2021	FSR and BER	DT and RF classifier	99.965%	0.998	**19**
[61]	2020	XGBoost	DT	90.85%	0.984	**19**
[90]	2020	A combination of PSO, grey wolf optimizer (GWO), firefly optimization (FFA), and GA	J48 and SVM	90.119%	0.969	**30**
[91]	2020	FFA–ant lion algorithm (ALO)	SVM, KNN, DT and NB	99.127%	0,935	**15**
[58]	2019	CFS and PSO	k-NN, SVM, and Naïve Bayes	92.877%	0.929	**13**
[57]	2019	RFE and RF for features selection	LR, SVM, NB, DT (C5.0), and Gradient Boost Machine	82.11%	0.86	**5**
[87]	2019	A combination of PSO, GA, and ant colony algorithm with REPT classifier for features selection	rotation forest and bagging classifiers	91.27%	0.913	**19**
[59]	2020	GR	Multi-layer perceptron Neural Network	76.96%	0.685	**30**
[56]	2018	AdaBoost	a combination of DT, NB, and ANN (Ensemble learning)	98.97%	0.979	**13**

Conclusions

- Most proposed IDSs suffer from a low detection accuracy or fail to detect all attacks
- Most proposed IDSs use a high number of features with a high computation cost
- RQA, a non-linear statistical analysis technique that uses the concepts of phase plane trajectory
- RQA shows its ability to identify hidden features that might not be observed using a single dimension system
- RQA improves detection accuracy and detect different types of attacks using a single feature
- The proposed approach outperforms most of the previous works in terms of accuracy and sensitivity

Computers & Security 122 (2022) 102906

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

A new intrusion detection system based on using non-linear statistical analysis and features selection techniques

Aliaa Al-Bakaa, Bahaa Al-Musawi*

Department of Electronic and Communications, Faculty of Engineering, University of Kufa, An Najaf, Iraq

ARTICLE INFO

Article history:
Received 30 January 2022
Revised 9 August 2022
Accepted 30 August 2022
Available online 5 September 2022

Keywords:
Intrusion detection system
Network threats
IoT security
UNSW-NB15 dataset
Recurrence quantification analysis
Machine learning algorithms

ABSTRACT

The increase in the number of connected devices to the Internet and Internet of Things (IoT) development accompanied a massive increase in the number and types of attacks. Most IoT devices have security vulnerabilities due to their limited computing and storage capabilities and specific protocols. Thus, there is essential to build Intrusion Detection Systems (IDSs) that can detect these threats that affect smart applications. Researchers examined different data mining techniques, statistical analysis techniques, and machine learning (ML) algorithms. In this paper, we propose a novel approach for building an anomaly-based intrusion detection system based on using a non-linear statistical analysis technique called recurrence quantification analysis (RQA). Our approach uses RQA to identify abnormal behavior in an individual feature extracted from a series of packets rather than inspecting the packet's payload. The proposed procedure implies finding the minimum number of features, applying the RQA technique to each effective feature separately, and applying different ML algorithms to classify the RQA measurements resulting from each effective feature. The system's performance was evaluated based on the accuracy and F-score using the UNSW-NB15 dataset. Results show our proposed approach's effectiveness in discovering hidden characteristics in the underlying series of an individual feature that leads to identifying different attacks. Besides, the proposed approach outperforms most previous works using one feature.