

EMPIRICAL VALIDATION OF STOCHASTIC APPROACHES TO DEVICE AND NETWORK STATE IDENTIFICATION

SWATI SHARMA



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI**

NOVEMBER, 2016

©Indian Institute of Technology Delhi - 2016
All rights reserved.

EMPIRICAL VALIDATION OF STOCHASTIC APPROACHES TO DEVICE AND NETWORK STATE IDENTIFICATION

by

SWATI SHARMA

Department of Computer Science and Engineering

Submitted

in fulfillment of the requirements of the degree of
Doctor of Philosophy

to the



Indian Institute of Technology Delhi

NOVEMBER, 2016

Certificate

This is to certify that the thesis titled **Empirical Validation of Stochastic Approaches to Device and Network State Identification** being submitted by **Swati Sharma** for the award of **Doctor of Philosophy in Computer Science & Engg.** is a record of bona fide work carried out by her under our guidance and supervision at the **Department of Computer Science & Engineering, Indian Institute of Technology Delhi**. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma.

Huzur Saran
Professor
Dept. of Computer Science & Engg.
Indian Institute of Technology Delhi.
Delhi, India.

Alefiya Hussain
Research Scientist
Information Sciences Institute
University of Southern California
Los Angeles, USA

Acknowledgements

Working towards my graduate thesis has been an incredible learning experience, both personally and professionally. I am grateful for His Grace and Guidance that made it possible. With the successful completion of this thesis, I would like to express my gratitude to all who made this journey possible.

I would like to thank my supervisor Prof. Huzur Saran for providing me with an opportunity to work under his guidance and introducing me to the world of research. His consistent support, encouragement and endless patience during this period taught me how to tackle hurdles in one's stride in an innovative manner and learn from past mistakes. He inspired me to identify my weaknesses and transform them to my strengths. I thank him for providing me with a welcome, secure, state-of-art working environment, and continuously guiding my passions in a gentle but persistent manner. It has made this journey interesting to undertake, even in the challenging moments. It has been my pleasure to learn from the best. I dedicate this thesis to him.

I express my heart-felt gratitude to my co-supervisor Dr. Alefiya Hussain for accepting to co-mentor me. I thank her for her belief in my abilities. I admire her patience in teaching me the benefits of careful planning in research, reminding me time and again the importance of minute details in experimental research. I was fortunate to pursue a semester internship with her at USC. This period helped me grow at a personal level and gave me an opportunity and platform to professionally interact with other distinguished researchers.

My sincere thanks to Dr. Praveen Bhagwat and Dr. Sohail Ahmed, Airtight Networks for providing me with an insight into research in industry and giving me exposure to wireless measurement research that paved the path for my dissertation. My special thanks to Dr. K.K. Ramakrishnan for enduring through my queries during the initial problem hunting phase of my work.

I express my gratitude to my student research committee members - Prof. B. N. Jain, Prof. Sanjiva Prasad and Prof. Brejesh Lall for constantly assessing my progress during the course of my research, identifying weaknesses in my work and suggesting ways to overcome them. I thank faculty members Dr. Sorav Bansal, Dr. Vinay Ribeiro, Dr. Parag Singla and Prof. Kolin Paul for brainstorming sessions. I thank Rajesh Kumar, S. S. Negi and Suresh Kumar for their timely assistance in all lab related issues. I also thank the department for providing

me with a comfortable work environment and other faculty and staff in the department for their constructive criticisms and support. Thanks to my lab friends Rakhi Tripathi, Mrinal Kumar, Nitesh Mor, Manoj Gupta, Namita Sharma, Nisha Jain, Smruti Padhy, Sakshi Tiwari for their great company.

I thank my loving family - my parents, my sister, my grandmother, my husband and my in-laws for their blessings and support during this period. I am grateful to my mother for her unconditional support and for silently teaching me by her actions that there are no parallels for hard work. Finally, a very special thanks to my father for always encouraging me to face all my demons head-on and teaching me that all enjoyment lies in the journey and not the end. I am fortunate to have him as my life-long friend and mentor. I will forever be grateful to him for sharing his personal and professional life experiences with me to help shape mine. I dedicate this thesis to him.

Swati Sharma

Abstract

The research work discussed in this thesis is centered on *robustly fingerprinting inherent behaviors in stochastic systems*. We discuss two research problems in this thesis. First, we focus on unique device identification based on inherent clock based fingerprints. Modern networks comprise of a mix of heterogeneous devices. So, any device identification approach must be able to uniquely identify them based on the same identification parameter and under varying host configurations, measurements and environment conditions. In other words, we emphasize on an empirical validation of the unique network device identification based on these clock behavior fingerprints under the influence of stochastic environments. Secondly, we work on fingerprinting (or characterizing) the inherent behavior of stochastic network experiments captured in their structure (that is, their executions on the network). These experiments typically are comprised of repeated executions of a mix of network protocols in order to optimize them or to create new ones. This fingerprinting/characterization can lead to verification of experiment executions and by extension, evaluating the success in repeated experiment executions (by comparing the fingerprints/characterizations). Thus, in this thesis, we propose data articulation and data aggregation based approaches for extensive empirical validation of network entities (network devices and network services/states/experiments). These two problems are not tightly coupled but instead are closely coupled. Our results are supported by hypothesis testing and analysis through ANOVA, KS, chi-square and correlation tests. These research problems are discussed briefly in the following text.

Modern enterprise networks contain heterogeneous devices. Last-mile connectivity to them is typically provided by branching the wired networks into wireless networks. This network transition relaxes the strict security prevalent in traditional wired networks. This transition, further, has liquidated the enterprise boundaries from desktops (in the wired network) to personal handhelds. As a result, the handheld devices contain both user's personal financial data as well as enterprise's confidential data. At the same time, network sessions have now become longer to accommodate frequent disconnections in the wireless medium. In case of a disconnection, the device connects back to its original session very conveniently in order to avoid a lengthy authentication process. Additionally, networks today implement a two-factor authentication process. Identification at the client end is typically performed by software-based parameters

(cookies, OS, browser, etc.). Most of the existing hardware based identification techniques require special hardware installation for their operation. If obtaining access to the device is possible, then it is also feasible to acquire a device image (disk image or flash memory image) and port it to either an emulated/virtualized platform or a different physical hardware altogether. In such a case, the operating system, session cookies and other software-parameters for device identification would remain the same.

Thus, if the device identification parameter is innate to the device hardware, such device cloning to an emulated hardware or a different physical hardware may be detected. This would not have been noticed at the software level. A fascinating alternative here is to complement the user authentication process with a stable, reliable hardware-dependent parameter. Any change in device hardware will thus alert the network administrator of a possible security breach. The network administrator can thereby add additional check(s) in server-end authentication to clear this security flag. Applications like forensics, attack attribution and IDS monitoring will benefit a lot by this alternative authentication process in the network.

After an empirical validation of network device identification based on characteristics of the device clock, our second goal is to fingerprint the inherent behavior of stochastic experiments. The experimental nature of networking research involves novel experiments for creation, optimization, design and execution of networking protocols. As a result, these experiments are repeated multiple times (not necessarily consecutively) on different network platforms (wired/wireless/emulation testbeds/simulators), network topologies and hardware apparatus for obtaining optimal results. Other researchers also re-execute these experiments to account for the human element of error. An experiment consists of three main components, namely, (a) deterministic (basic programming code targeted to run a certain way), (b) non-deterministic (rare error cases in programming code or dynamic network behavior during experiment execution), and (c) opportunistic (cyber-attack models following multiple code paths for successful execution). Further, these experiments are complex and stochastic in nature. Diverse sources of variation in an experiment render attaining repeatability and verifiability very difficult. For instance, an experiment's (1) physical apparatus, (2) topology, (3) software code or binaries, (4) input parameters, (5) hardware and software configuration of the nodes, (6) procedure, (7) measurement and analysis process, (8) output, (9) cross-traffic involved, (10) other network constraints originating from the dynamic network behavior, etc.

A user repeating his own experiment has a decent idea about its (experiment's) execution and may identify if the experiment is not executed accurately. But another researcher may not possess that level of understanding into other researcher's work. Thus, the ability to accurately repeat an experiment and verify its accuracy is critical. It would be fascinating to tackle this research problem just like the network device identification problem described above. In other words, if a network experiment could be identified by an identification mechanism (that

is a characterization or a signature) based on its execution details then this information could be used in the future to verify repeated executions. At the fundamental level, this information could be the number of events, type of events and the sequence of occurrence of these events taking place in the experiment execution. A repository or fingerprint database could be maintained for storing the signatures/characterizations of different network experiments carried out by the researcher. Verifying an experiment's correct execution will then be reduced to the trivial task of comparing characterizations of fresh experiment executions with the experiment characterizations stored in the repository.

The main contributions of this thesis are the following.

- In order to provide a reliable means of hardware-based network-device identification, we study and compare the existing device ID approaches. We then select clock-skew based device identification for further empirical validation. This idea was presented (or introduced) by Kohno et. al. We emphasize on the local fingerprinting in a network (providing last mile access to client devices) and monitoring of active network hosts. No such work existed when we began this research in 2010. The rest are subsequent works. The contribution of our work is that we take a closer look at this approach and provide a *systematic empirical validation of clock skew fingerprint stability* by performing a *multi-dimensional comparison (3X3X4) of factors* influencing the clock skews for all heterogeneous devices. These dimensions are timestamp measurement methods, target host measurement environments, and target host configurations. These variables in measurement are promising candidates that intuitively would be likely to cause differences in clock skew. Thus, in this research, we ask questions like examining if the approach worked, how well it worked and under what conditions did it work?
- We provide *optimizations for the clock-skew based device identification approach* and identify the factors that affect a device's skew estimate in a heterogeneous network. Our research includes a thorough experimental validation of skew behavior for smart handheld devices. No such work existed when we began this research in 2010. The rest are subsequent works. We modify the optimized clock skew estimation technique to tailor it according to handheld device sleep routines to avoid latency in captured timestamp value packets on the network. We *discover skew jumps in mobile handheld devices, induced by varying operating power modes and ambient device temperatures*. We apply the optimized identification approach to other devices in a heterogeneous network, namely, virtual machines, testbed nodes. We demonstrate that *VM skews show a completely different skew behavior as compared to that published by Kohno et. al.* We also demonstrate that *testbed node skew behavior* is unexpectedly different from the clock skew trends

of both physical and virtual machines. We finally, leverage the optimized skew based identification approach to distinguish between individual cores of a multi-core machine.

- Once a characterization of skew based identification in modern heterogeneous networks has been evaluated, we now examine the characterization of stochastic network services (experiments). Our framework directly characterizes a network experiment's execution. The aforementioned characterization facilitates the comparison of repeated experiment executions based on event logs and measurement data. This framework results in the creation of Time-Series and Event Based Markov Chain characterizations. Comparison of these characterizations is performed by standard distance measures like Kullback-Leibler Divergence, Total Variation Distance and Euclidean Distance.
- We demonstrate that our framework is statistically rigorous, sound and sensitive to changes in experiment configuration like hardware, topology, traffic and mobility. We extensively test our framework on a few experiment classes (opportunistic and heavy-tailed) over different experimentation platforms (simulators, emulated testbeds and real-networks). To do this, we create multiple configuration-based variations of these experiment classes on these experimentation platforms. Subsequently, we perform hundreds of executions for each variation and then compare these executions with our framework. We also establish that these characterizations are feasible under transformations of scale and complexity.

With the proposed device identification strategy, we tracked 152 devices, 17 handhelds, 48 virtual machines and 22 testbed nodes for a period of 9 months each. Clock skew estimate for every device was found to be stable under almost all investigation parameters. Significant differences between skew behavior for desktops and handhelds were observed. Namely, desktop skew estimates were affected by NTP updates while handheld skew estimates were affected by varying operating power modes and ambient device temperatures. Maximum value for error threshold in the skew estimates for all devices was 0.3 ppm. We found that to achieve the minimum error threshold and the most stable skew estimate, a minimum of 70 packets were required to calculate one skew estimate. Variable skew jumps were observed in handhelds with changing operating power states (AC power/battery operated) and varying ambient device temperatures (-18 to 45°Celsius). We also demonstrated that these jumps were dependent on the type & manufacturer of the device. VM skews were found to be neither inconsistent nor extremely large as outlined in previous published research. We also showed that testbed nodes exhibited similar skew estimates upto the 3rd decimal place. Thus, we found that for a moderate-size network, clock skew based fingerprints provided a reasonable mode of identification under a variety of host, configuration and measurement environments.

With the proposed network service/state characterization strategy, we analyzed 1500 DETER testbed DNS-cache poisoning attack experiment trials, 3200 ns2 webtraffic-generation experiment trials, 1000 real-time webtraffic-generation experiment trials and 300 ns3 wireless mobile traffic-generation experiment trials. Our framework based experiment characterizations were found to be sensitive to experiment configuration and as a result the impact of topology, traffic, hardware and mobility on these characterizations was investigated. We demonstrated that our framework was statistically rigorous, sound and sensitive to changes in the experiment's configuration (hardware, topology, traffic and mobility). We extensively tested our framework by performing hundreds of experiment executions on a few experiment classes (opportunistic and heavy-tailed experiments) and heterogeneous experimentation platforms (simulators, emulated testbeds and real-networks). Thus, we found that our framework provided a direct and precise method to compare two executions of a stochastic networking experiment with Markovian dependencies for simulations, emulation-based testbeds and real-time network experiments.

Thus, the research in this thesis provides an empirical study to help the reader to understand the full-scale capabilities of skew based fingerprint in a heterogeneous network under a set of varying configurations, measurement and environment conditions. This research also lays the foundation for generating a validity management framework for determining verifiability and repeatability in experiment executions.

Contents

Acknowledgements	7
Abstract	9
List of Figures	vii
List of Tables	xi
1 Background and Motivation	1
1.1 Background for Network Device Identification	2
1.1.1 Biometric Fingerprints	3
1.1.2 Host/Device Identification Sub-Classification	3
1.1.3 Optimal Device Fingerprints	4
1.2 Background for Network Services Identification	5
1.2.1 What is a Network Experiment?	5
1.2.2 Stochastic Nature of Experiments	5
1.2.3 Need for Characterization in Networking Experiments	6
1.2.4 Need for Repeatability and Verifiability in Networking Experiments	6
1.2.5 Challenges Encountered by Testbed Users	7
1.2.6 Possible Models for Experiment Characterization	8
1.3 Background on Experimental Measurement Research	8
1.3.1 Sample Size and Significance Level Determination	10
1.3.2 Logical Steps of Scientific Experimentation	10
1.4 Motivation for Thesis Research	13
1.4.1 Motivation - Device Identification in Modern Heterogeneous Enterprise Networks	13
1.4.2 Motivation - Identification and Verification of Network Services/Exper- iments	13
1.5 Application Domains for Identification of Network Entities	14

1.5.1	Device Identification in Modern Heterogeneous Enterprise Networks	14
1.5.2	Identification and Verification of Network Services/Experiments	16
1.6	Thesis Contributions	17
1.7	Thesis Overview	19
1.7.1	Chapter 2 : Unique Device Identification in a Heterogeneous Network	19
1.7.2	Chapter 3 : Clock-Skew based Device Identification in Modern Heterogeneous Networks	19
1.7.3	Chapter 4 : Device Fingerprinting beyond Physical Machines	19
1.7.4	Chapter 5 : Experiment identification and Comparison of Experiment Runs.	20
1.7.5	Chapter 6 : Conclusion and Future Work	20
2	Device Identification in a Heterogeneous Network	21
2.1	Significance of Device Fingerprinting	21
2.1.1	Do we need it?	22
2.1.2	What are the existing options?	23
2.1.3	Client Classification for the Network Administrator	24
2.2	Fingerprinting Technique Classification	25
2.2.1	Subject of Identification	25
2.2.2	Operation Mechanism	25
2.2.3	Hardware/Software Configuration Parameters	26
2.2.4	Layer of Operation	27
2.3	Existing Approaches to Unique Device Identification	29
2.3.1	Session Cookie based fingerprinting	29
2.3.2	Physical layer characteristics based fingerprinting	31
2.3.3	OS fingerprinting	31
2.3.4	Sequence number anomalies based fingerprints	32
2.3.5	Clock skew based device identification	32
2.3.6	Power management handling and random back-off interval calculation based fingerprints	33
2.3.7	Analog signal fingerprinting using the matched filter approach	33
2.3.8	Spectral analysis for coding rate sequence and data transmission control based fingerprints	33
2.3.9	Active behavioral fingerprinting	34
2.3.10	Coupling OS, wireless NIC driver and machine characteristics	34
2.3.11	Acknowledgment frame delay fingerprinting	35
2.3.12	Browser fingerprints	35

2.3.13	Acoustic fingerprints for handhelds	35
2.3.14	Network Device Identification	36
2.4	Why Clock Skew Fingerprints?	36
2.5	Clock Basics	37
2.5.1	Clock Sources	37
2.5.2	Clock Synchronization over the network	39
2.5.3	Timestamp Sources	39
3	Clock-Skew Based Device Fingerprinting in Heterogeneous Networks	43
3.1	Clock Skew Based Fingerprints	44
3.1.1	Key Idea - Clocks Shift in Time	44
3.1.2	Clock Terminology Used	45
3.1.3	Clock Skew Estimation	45
3.1.4	Interpretation of Observations	47
3.1.5	Previous work on clock skew fingerprints	48
3.1.6	Applications	49
3.1.7	Contributions	49
3.1.8	Potential Countermeasures	50
3.2	Related Work	50
3.3	Experimental Setup	52
3.4	Measurement Methodology	55
3.4.1	TCP Mode	55
3.4.2	Batch & Continuous ICMP Modes	56
3.4.3	Tailoring Skew Extraction for Smart Devices	58
3.4.4	Complexity	58
3.5	Clock Skew Stability	59
3.5.1	Across Measurement Methodologies	59
3.5.2	Across Target Device Environments	63
3.5.3	Across Target Device Configurations	66
3.6	Effects of Hardware Aging	68
3.6.1	Experiment Setup and Measurement Methodology	69
3.6.2	Data and Observations	69
3.6.3	Analysis and Results	69
3.7	Device Tracking Sensitivity	70
3.8	Optimizations to the original design	72
3.9	Future Work and Limitations	73
3.10	Conclusion	75

4	Clock-Skew Fingerprints Beyond Physical Machines	77
4.1	Introduction	77
4.1.1	Contributions	78
4.2	Related Work	78
4.3	Experiment Setup	79
4.3.1	Handhelds	79
4.3.2	Virtual Machines	80
4.3.3	Testbed Nodes	80
4.3.4	Multi-core Machines	81
4.4	Handheld Skews	81
4.4.1	Data	82
4.4.2	Analysis	82
4.4.3	Voltage-Temperature Behavior	84
4.5	Virtual Machine Skews	86
4.5.1	Data and Observations	87
4.5.2	Analysis	87
4.6	Testbed Skews	87
4.6.1	Data and Observations	88
4.6.2	Analysis	88
4.7	Multi-Core Skews	89
4.7.1	Data and Methods	89
4.7.2	Observations and Analysis	90
4.8	Evaluation of Results and Discussion	95
4.8.1	Statistical Analysis and Hypothesis Testing	95
4.8.2	Validation of Processor Clock Drift	96
4.8.3	Modified Skew Fingerprint Strategy	97
4.8.4	Client Device Skew Fingerprint in Modern Networks	97
4.9	Future Work & Limitations	98
4.10	Conclusion	99
5	Towards Repeatability & Verifiability in Networking Experiments: A stochastic framework	101
5.1	Introduction	102
5.1.1	Denotations	102
5.1.2	Concept Illustration with File Transfer Experiment	103
5.1.3	Contributions	103
5.1.4	Applications - Is This Problem Worth Solving?	104

5.2	Related Work	104
5.3	Key Idea	106
5.3.1	Brief Summary of Approach	108
5.4	Framework Design	108
5.4.1	Experiment Characterization Formulation	108
5.4.2	Measures of Similarity (or Agreement)	109
5.4.3	Why Markov Model-based framework	109
5.4.4	Basic Design Choices	111
5.5	Framework Selection - Which Framework to Use?	111
5.6	Framework Creation - Time-Series Based Characterizations	112
5.6.1	Algorithm Pseudo-Code	113
5.6.2	Is Time-Series Based Characterization enough?	114
5.7	Framework Creation - Event Based Markov Chain Characterizations	114
5.7.1	Markov Chain Creation	114
5.7.2	Accuracy	114
5.7.3	Algorithm Pseudo-Code	115
5.7.4	State Identification	116
5.7.5	Granularity for State Definition	116
5.7.6	Packet Logging Time in State Identification	117
5.7.7	Packet Sequence defines State	117
5.7.8	Multiple hosts in Experiment Execution	118
5.7.9	Different Packet Routes in State Identification	118
5.7.10	Aggregate and Individual Hops	119
5.7.11	Occurrence of Rare Events	119
5.8	Framework Comparison - Similarity between Experiment Characterizations	119
5.8.1	Characterization Comparison Aspects	120
5.9	Simple Markov Chain Creation Sample	121
5.10	Complex Markov Chain Creation Sample	123
5.10.1	Experiment Background	123
5.10.2	Experiment Execution	124
5.10.3	Experiment Characterization	125
5.11	Experiment Setup & Data Collection	128
5.11.1	DNS Cache Poisoning Attack Experiment	129
5.11.2	Web-Traffic Generation Simulations	131
5.11.3	Wireless Mobility Simulations	133
5.12	Observations & Analysis	134
5.12.1	Time-Series Based Experiment Characterizations	134

5.12.2	Need for Event-Based Markov Chain Characterizations	136
5.12.3	Analysis of Measurements : Markov Chain Characterizations	137
5.13	Evaluation of Results	141
5.13.1	Statistical Analysis & Hypothesis Testing	142
5.13.2	1-step and 2-step Markov Model Comparison	143
5.13.3	Computational & Error Complexity	145
5.13.4	Metrics for Experiment Trial Comparison	146
5.14	Future Work and Limitations	146
5.15	Conclusions	147
6	Conclusion and Future Work	149
6.1	Summary of Contributions	149
6.2	Limitations and Future Directions	151
6.2.1	Device Identification in Modern Heterogeneous Enterprise Networks	151
6.2.2	Identification and Verification of Network Services/Experiments	153
	Bibliography	155
	List of Publications	173
	Biography	175

List of Figures

1.1	Illustration of ideal characteristics of an optimal device fingerprint technique.	5
1.2	Schematic summarizing the generation of experiment characterizations/signatures.	9
1.3	Schematic summarizing the generation of experiment characterizations/signatures.	9
1.4	Illustration of Logical Steps of Scientific Experimentation process [1].	11
2.1	Illustration of the device fingerprinting concept.	22
2.2	Illustration of stages in device fingerprinting.	23
2.3	Sub-classification for clients on a network.	24
2.4	Sub-classification for fingerprinting techniques on a network.	26
2.5	Breakup of fingerprinting techniques based on parameter of identification used for Radio Frequency Fingerprinting.	27
2.6	Breakup of fingerprinting techniques based on parameter of identification used for Data Link Layer Fingerprinting.	28
2.7	Breakup of fingerprinting techniques based on parameter of identification used for Clock Skew Fingerprinting.	28
3.1	Illustration of time shifts in clocks.	44
3.2	Interpretation of skew behavior from measurement point and target host timestamps collected from ICMP timestamp response packet headers for a single device.	47
3.3	Distinctive Skew behavior for 16 homogeneous target devices.	48
3.4	Network layout of the experimental setup illustrating multiple measurement points and a diverse set of target devices.	54
3.5	Illustration of TCP timestamp collection mode in brief.	55
3.6	Illustration of batch ICMP timestamp collection mode in brief.	56
3.7	Illustration of continuous ICMP timestamp collection mode in brief.	57

3.8	Optimized ICMP Timestamp Request-Response Packet Sequence for Handheld Timestamp Extraction.	58
3.9	Error Threshold and Skew Estimate variation with respect to number of packets utilized for skew computation.	63
3.10	Scatter plot for illustration of skew estimates with respect to the device identifications.	71
4.1	Typical experimental network layout for traditional clock skew identification as compared to clock skew extraction layout for processor cores.	81
4.2	'rdtsc' value progression for dual cores in xv6.	91
4.3	Non-uniform behavior of consecutive rdtsc value differences for OS-based logging in Ubuntu installed uni-processor system.	91
4.4	Non-uniform behavior of consecutive rdtsc differences for OS-based logging in Ubuntu installed multi-processor system.	93
4.5	Non-uniform behavior of consecutive rdtsc differences for kernel module based logging in Ubuntu installed multi processor system.	93
4.6	rdtsc behavior after disabling tickless property in Ubuntu installed multi processor system.	94
5.1	Depiction of Experiment Trial Comparison.	106
5.2	Schematic representing testing the match between current & registered experiment trials.	107
5.3	Topology Illustration for 2 clients instead of 1 for the sample Markov chain creation example.	118
5.4	Topology Illustration for 1 client instead of 2 in a client-server model used for the sample Markov chain creation example.	121
5.5	Transition Diagram for the sample Markov chain creation example.	122
5.6	Comparison graph for intra-configuration and inter-configuration comparisons.	123
5.7	Illustration of basic DNS Kaminsky cache poisoning attack experiment topology.	125
5.8	Illustration of the possible paths for DNS Kaminsky cache poisoning attack.	127
5.9	Transition Diagram for the sample Markov chain creation.	128
5.10	Illustration of the modified topology for the basic DNS Kaminsky cache poisoning attack experiment.	130
5.11	Illustration of the dumbbell topology for the web-traffic generation experiment.	132
5.12	Illustration of the modified dumbbell topology for the web-traffic generation experiment.	132
5.13	CDF based characterization of inter-arrival time series from trials.	135

5.14	Time series from 3 trials. Plot I and II have identical experiment configurations and hence "similar" behaviors captured through time series. Changing configuration results in different behavior in Plot III.	137
5.15	Distance measures for Markov Chain characterizations - DNS cache poisoning.	139
5.16	Distance measures for Markov Chain characterizations - web-traffic generation.	139
5.17	Distance measure comparisons for Markov Chain characterizations of wireless mobility experiment.	140
5.18	Distance measure comparisons for 1-step Markov Chain characterizations of DNS cache poisoning experiment.	144
5.19	Distance measure comparisons for 2-step Markov Chain characterizations of DNS cache poisoning experiment.	144

List of Tables

1.1	Summary of Experimental Platform Conditions from [2]	7
2.1	A summary of existing alternative fingerprint techniques.	30
3.1	Experimental Configuration for Clock Skew Measurement Study at IIT Delhi.	53
3.2	Target Device Identification Details	61
3.3	Measurement Interval Dependent Skew Scatter.	62
3.4	Clock Skew stability across different investigation parameters of system times-tamp collection methodology.	63
3.5	Measurement Point Dependent Skew Behavior.	65
3.6	Clock Skew stability across different investigation parameters of target host environments.	66
3.7	Power State Dependent Handheld Skew Behavior	67
3.8	Target Device OS Dependent Skew Behavior.	67
3.9	Clock Skew stability across different investigation parameters of target host configuration.	68
3.10	Device identification details for the target device subset illustrated in figure 3.10.	70
4.1	Ambient Environment Temperature Dependent Skew Behavior for devices running on AC Power	83
4.2	Ambient Environment Temperature Dependent Skew Behavior for devices running on Battery Power	83
4.3	Ambient Environment Temperature Dependent Voltage Behavior for S3 running on Battery Power at -18°C.	85
4.4	Ambient Environment Temperature Dependent Voltage Behavior for S3 running on AC Power at -18°C.	85
4.5	Ambient Environment Temperature Dependent Voltage Behavior for S3 running on Battery Power at 45°C.	85

4.6	Ambient Environment Temperature Dependent Voltage Behavior for S3 running on AC Power at 45°C.	86
4.7	Virtual Machine Clock Skews	89
5.1	Emulation Experiments (DNS Kaminsky Cache Positioning Attack): Left table assigns identifier to experiment designs, right table assigns identifier to comparisons of designs from the left.	129
5.2	Simulation Experiments (Web Traffic Generation): Left table assigns identifier to experiment designs, right table assigns identifier to combinations from the left.	131
5.3	Wireless Mobility Experiments: Left table assigns identifier to experiment designs, right table assigns identifier to combinations from the left.	133
5.4	Correlation coefficient comparisons for trials. Values close to 1 indicate strong correlation.	136
5.5	ANOVA test values for trial comparisons.	143