

# Network Behavior Analysis

Kuai Xu

# Network Behavior Analysis

Measurement, Models, and Applications

Kuai Xu  
School of Mathematical  
and Natural Sciences  
Arizona State University  
Glendale, AZ, USA

ISBN 978-981-16-8324-4      ISBN 978-981-16-8325-1 (eBook)  
<https://doi.org/10.1007/978-981-16-8325-1>

© Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.  
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Preface

As the Internet continues to grow in size and complexity, the challenge of effectively provisioning, managing, and securing it has become inextricably linked to a deep understanding of network behaviors of networked systems and Internet applications. While there exists an extensive body of research publications on traffic classifications, Internet measurement, network security, and digital forensics, there are few books dedicated to network behavior analysis. This book provides a comprehensive overview of network behavior analysis that focuses on the study of network traffic data to provide critical insights into the behavioral patterns of networked systems such as servers, desktops, smartphones, and the Internet of Thing (IoT) devices and Internet applications such as web browsing, electronic mails, file transfers, online gaming, video streaming, and social networking. The objective of this book is to fill the book publication gap in network behavior analysis which has recently become an increasingly important component of comprehensive network monitoring and security solutions for backbone networks, enterprise networks, data center networks, home networks, and emerging networks such as 5G networks, vehicle networks, and IoT networks.

Network behavior analysis is an end-to-end process of collecting, extracting, analyzing, modeling, and interpreting network behavior of end systems and network application from Internet traffic data such as TCP/IP data packets and network flows. This book presents the fundamental principles and best practices for network behavior analysis. Relying on data mining, machine learning, information theory, probabilistic graphical and structural modeling, this book explains what, who, where, when, and why of communication patterns and network behavior of networked systems and Internet applications. The book also discusses the benefits of network behavior analysis for the applications of cybersecurity monitoring, Internet traffic profiling, anomaly traffic detection, and emerging application detection.

This book is of particular interest to researchers and practitioners in the fields of Internet measurement, traffic analysis, cybersecurity since this book brings a spectrum of innovative techniques to develop behavior models, structural models, graphic models of Internet traffic and presents how to leverage these results from

these models for a broad range of real-world applications in network management, security operations, and cyber intelligent analysis.

The major benefits of reading this book include (1) learning the principles and practices of measuring, modeling, and analyzing network behavior from massive Internet traffic data; (2) making sense of network behavior for a spectrum of applications ranging from cybersecurity, network monitoring and emerging application detection; and (3) understanding how to explore network behavior analysis to complement traditional perimeter-based firewall and intrusion detection systems to detect unusual traffic patterns or zero-day security threats via data mining and machine learning techniques. The prerequisite for reading this book is a basic understanding on TCP/IP protocols, data packets, network flows, and Internet applications.

Phoenix, AZ, USA  
October 2021

Kuai Xu

# Acknowledgements

I would like to acknowledge the contributions to the research publications, which build the theoretical and system foundations of this book, from my collaborators and co-authors: Supratik Bhattacharyya, Jianhua Gao, Lin Gu, Yaohui Jin, Yinxin Wan, Feng Wang, Guoliang Xue, and Zhi-Li Zhang.

I would also like to thank Sprint, Center for Applied Internet Data Analysis (CAIDA) based at the University of California's San Diego Supercomputer Center, and University of Oregon Route Views Project for making Internet traffic datasets and routing table datasets available to the research projects in the book.

I am grateful for the constructive feedbacks and insightful comments from the reviewers, which have greatly improved the technical presentation of the book. I also would like to thank Springer editorial staff for their support and encouragement from the beginning of the book project to the final publication.

The research projects in this book are partially supported by three research grants (CNS-1218212, DMS-1737861, and CNS-1816995) from US National Science Foundation (NSF).

# Contents

- 1 Introduction** ..... 1
  - 1.1 What is Network Behavior Analysis ..... 1
  - 1.2 Network Behavior Measurement and Modeling ..... 2
  - 1.3 Benefits of Network Behavior Analysis ..... 3
  - 1.4 Book Overview and Organization ..... 4
  - References ..... 4
- 2 Background of Network Behavior Analysis** ..... 7
  - 2.1 Internet Measurement and Analysis ..... 7
  - 2.2 Data Collection for Network Behavior Analysis ..... 9
  - 2.3 Preliminaries of Network Behavior Analysis ..... 11
    - 2.3.1 Information Theory and Entropy ..... 11
    - 2.3.2 Graphical Analysis ..... 13
  - References ..... 15
- 3 Behavior Modeling of Network Traffic** ..... 21
  - 3.1 Behavior-Oriented Network Traffic Modeling ..... 21
    - 3.1.1 What is Network Behavior ..... 21
    - 3.1.2 Traffic Features in Network Behavior ..... 22
    - 3.1.3 Behavioral Entities ..... 22
    - 3.1.4 Real-World Network Traffic Datasets ..... 23
  - 3.2 Identifying Significant Behavioral Entities ..... 24
    - 3.2.1 Significant Behavioral Entities ..... 24
    - 3.2.2 Adaptive Thresholding Algorithm ..... 25
    - 3.2.3 Extracting Significant Traffic Clusters ..... 26
  - 3.3 Network Behavior Modeling ..... 29
    - 3.3.1 Network Behavior Modeling ..... 29
    - 3.3.2 Network Behavior Classifications ..... 30

3.4	Network Behavior Dynamics .....	32
3.4.1	Temporal Properties of Behavior Classes .....	32
3.4.2	Behavior Dynamics of Individual Clusters .....	33
3.5	Summary .....	37
	References .....	37
<b>4</b>	<b>Structural Modeling of Network Traffic .....</b>	<b>39</b>
4.1	Communication Structure Analysis .....	39
4.1.1	Dominant State Analysis .....	39
4.1.2	Communication Structure of Networked Systems and Internet Applications .....	41
4.2	Exploring More Traffic Features .....	44
4.3	Summary .....	47
	References .....	48
<b>5</b>	<b>Graphical Modeling of Network Traffic .....</b>	<b>49</b>
5.1	Cluster-Aware Network Behavior Analysis .....	49
5.2	Modeling Host Communications with Bipartite Graphs and One-Mode Projections .....	50
5.3	Similarity Matrices and Clustering Coefficient of One-Mode Projection Graphs .....	51
5.3.1	Similarity Matrices .....	51
5.3.2	Clustering Coefficients .....	53
5.4	Discovering Behavior Clusters via Clustering Algorithms .....	54
5.4.1	Partitioning Similarity Matrix with Spectral Clustering Algorithm .....	54
5.4.2	Clustering Analysis of Internet Applications .....	57
5.5	Traffic Characteristics and Similarity of Behavior Clusters .....	59
5.5.1	Making Sense of End-Host Behavior Clusters .....	59
5.5.2	Distinct Traffic Characteristics of Behavior Clusters .....	63
5.5.3	Exploring Similarity of Internet Applications .....	66
5.6	Summary .....	68
	References .....	68
<b>6</b>	<b>Real-Time Network Behavior Analysis .....</b>	<b>71</b>
6.1	Real-Time Network Measurement and Monitoring .....	71
6.2	Real-Time System for Network Behavior Analysis .....	72
6.2.1	Design Guidelines .....	73
6.2.2	System Architecture .....	73
6.2.3	Key Implementation Details .....	75
6.3	Performance Evaluation .....	80
6.3.1	Benchmarking .....	80
6.3.2	Stress Test .....	84
6.4	Sampling and Filtering .....	86
6.4.1	Random Sampling .....	86
6.4.2	Profiling-Aware Filtering .....	87



6.5	Summary	91
	References	92
<b>7</b>	<b>Applications</b>	93
7.1	Profiling Internet Traffic	93
7.1.1	Server/Service Behavior Profiles	94
7.1.2	Heavy-Hitter Host Behavior Profiles	96
7.1.3	Scan/Exploit Profiles	96
7.1.4	Deviant or Rare Behaviors	99
7.2	Reducing Unwanted Traffic on the Internet	100
7.2.1	Unwanted Exploit Traffic on the Internet	101
7.2.2	Characteristics of Unwanted Exploit Traffic	103
7.2.3	Strategies of Reducing Unwanted Traffic	107
7.2.4	Sequential Behavior Analysis	108
7.3	Cluster-Aware Applications of Network Behavior Analysis	111
7.3.1	End-Host Network Behavior Clusters	112
7.3.2	Network Application Behavior Clusters	115
7.4	Summary	117
	References	117
<b>8</b>	<b>Research Frontiers of Network Behavior Analysis</b>	119
8.1	Network Behavior Analysis in the Cloud	119
8.1.1	Background	120
8.1.2	Profiling-as-a-Service in the Cloud	121
8.1.3	Architecture of Profiling-as-a-Service for Network Behavior Analysis	121
8.1.4	Designing the Profiling-as-a-Service Infrastructure	123
8.2	Network Behavior Analysis in Smart Homes	127
8.2.1	Background	128
8.2.2	Traffic Monitoring Platform for Home Networks	129
8.2.3	Characterizing Home Network Traffic	131
8.2.4	Unwanted Traffic Towards Home Networks	139
8.3	Network Behavior Analysis for Internet of Things	141
8.3.1	Background	142
8.3.2	IoT Traffic Measurement and Monitoring	143
8.3.3	An IoT Traffic Measurement Framework via Programmable Edge Routers	145
8.3.4	Multidimensional Behavioral Profiling of IoT Devices	146
8.3.5	Exploring the Applications of Multidimensional Behavioral Profiling	153
8.4	Summary	159
	References	159