PhD Thesis Approval Form

THESIS TITLE: Jurisdictional Arbitrage: Quantifying and Counteracting the Threat of Government Intelligence Agencies Against Tor

AUTHOR: Sarah L. Cortes

Ph.D. Thesis Approved to complete all degree requirements for the Ph.D. Degree in Computer Science- Information Assurance.

Thesis Advisor-Prof. Thomas H. Koenig  
Date: 9/27/17

Thesis Reader-Prof. László Barabási  
Date: 9/27/17

Thesis Reader-Prof. Quevvera Noubir  
Date: 9/27/17

Thesis Reader-Paul Syverson, Ph.D  
Date: 9/28/17

GRADUATE SCHOOL APPROVAL:

Director, Graduate School  
Date: 10/3/17

COPY RECEIVED IN GRADUATE SCHOOL OFFICE:  
Recipient's Signature  
Date: 10/3/17

Distribution: Once completed, this form should be scanned and attached to the front of the electronic dissertation document (page 1). An electronic version of the document can then be uploaded to the Northeastern University-UMI website.
Jurisdictional Arbitrage:
Quantifying and Counteracting the Threat of Government
Intelligence Agencies Against Tor

A dissertation presented in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the field of
Computer Science, Information Security and Assurance
by
Sarah Lewis Cortes

Committee:

Prof. Thomas H. Koenig- Professor, Interdisciplinary with the College of Computer and Information Science (CCIS) and the College of Social Sciences and Humanities (CSSH) at Northeastern University (Advisor)

Prof. László Barabási- Robert Gray Dodge Professor of Network Science, Distinguished University Professor, Interdisciplinary with the College of Computer and Information Science (CCIS) and the College of Science (COS) at Northeastern University

Prof. Guevara Noubir- Professor, Director of Cybersecurity and Information Assurance Graduate Program, College of Computer and Information Science (CCIS) at Northeastern University

Paul Syverson, Ph.D., United States Naval Research Laboratory, Center for High Assurance Computer Systems

Northeastern University
College of Computer and Information Science
College of Engineering
College of Social Sciences and Humanities
Boston, Massachusetts, USA
September, 2017

Copyright 2017 Sarah L. Cortes
Abstract

Recent events continue to expose the ability and commitment of Government Intelligence Agencies (GIAs) to conduct cross-border surveillance of the Internet. These revelations have significant consequences for anonymous communication systems like Tor. Current adversarial models do not incorporate international surveillance, meaning that realistic adversaries are much more powerful than what is assumed by prior work. In this work, we take the first steps towards quantifying the risk of surveillance posed by GIAs. We use legal and technical data to assess the hostility of each country to Internet traffic, and build a graph of the intelligence treaties between countries to identify cross-border surveillance capabilities. Based on this data, we develop metrics that quantify the ability of an adversarial GIA to conduct surveillance in any other country. We apply our risk metrics to the current state of the Tor anonymity network and discover that the majority of Tor users are at significant risk to passive deanonymization attacks by GIAs. We incorporate our metrics into alternative Tor relay selection algorithms, and show that the resulting circuits have significantly reduced surveillance risk, compared to Tor’s standard relay selection algorithm.
Acknowledgements
# Table of Contents

Thesis Approval Form  
Title Page  
**Abstract**  
4  
**Acknowledgements**  
7  
**List of Tables**  
7  
**List of Figures**  
7  
**List of Equations**  
7  
**Roadmap**  
8  
**Overview**  
9  
1 Introduction to Jurisdictional Arbitrage  
12  
2 Background  
14  
2.1 The Public Tor Network  
14  
Path Selection in Tor  
16  
Attacks Against Tor  
16  
AS and IXP-Level Attacks  
17  
2.2 Threat Model  
17  
3 Related Work  
18  
3.1 Computer Science: Network Security  
18  
3.2 Network Science: Communications and Lawful Intercept Cartels  
18  
3.3 Law: Privacy, 4th Amendment, Search and Seizure, Intern’l Criminal Treaties  
19  
4 Research Methodology and Results  
20  
4.1 Jurisdictional Arbitrage: Quantifying Threats to Anonymous Communication Networks by GIAs  
20  
Defining Hostility Factors and Collecting Hostility Factor Data  
20  
Internet Traffic Surveillance: “Lawful Intercept” (LI)  
22  
Internet Traffic Surveillance: Data Retention Laws  
23  
Internet Censorship- Takedown Requests and User Monitoring  
24  
Internet Censorship- Direct Blocking  
24  
Internet Traffic Surveillance: Int’l Law Enforcement Treaties (MLATs)  
25  
4.2 MLAT Database: Documenting and Classifying worldwide MLATs  
26  
4.3 Jurisdictional Arbitrage: Defining Metrics for Quantifying Risk from GIAs  
31
Country Hostility Metric 31
Connection Risk Metric 33
Circuit Risk Metric 34

4.4 Jurisdictional Arbitrage: Applying Risk Metrics to Countries and Circuits 35
  4.4.1 Evaluating Country Hostility 35
  4.4.2 Evaluating Circuit Risk 37
    Key Findings and Observations 37
    Circuit Dataset 39
    Distribution of Circuit Risk 39
    Impact of Source and Destination on Circuit Risk 40
  4.3.4 Impact of Circuit Risk 41
    Source Risk versus Guard Risk 42

4.5 Jurisdictional Arbitrage: Novel relay and path selection algorithms and overlays 43

5 Related Research Projects 44
  5.1 MLATs and the Undersea Cable Network 44
  5.2 MLATs and Legalizing Domestic Surveillance: The Role of Mutual Legal Assistance Treaties in Deanonymizing TorBrowser Technology 45
  5.3 MLAT.is: Web portal with network database query and visualization capabilities 45
  5.4 MLAT Graph: Measuring the worldwide MLATs country cartel graph 45
  5.5 Dynamic Triggering: MLATs in Lawful Intercept for communications networks 45
  5.6 Jurisdiction of the Darknet: Successful LEA De-Anonymization 46
  5.7 CircuitBlasTor: Measuring The Tor Graph for Randomness 47
  5.8 Diameter of the Darknet-hyperlink analysis 49

References 49

Appendix A- Contributions 62
  By Contribution to Knowledge 62
  By Research Artifacts 63
  By Scientific or Academic Field 63
  By Key Findings and Observations 64

Appendix B- Research Deployment and Administration 66

Appendix C- Committee Background and Justification 67

PhD Thesis Approval Form 71

Graphics 72
List of Tables
Table 1: Snapshot of our dataset for 10 countries with the most Tor relays
Table 2: Top 21 most hostile countries according to our metric

List of Figures
Figure 1. CDF of nonzero hostility scores Hc for ISO 3166-1 countries
Figure 2. Hostility of countries versus the number of Tor relays in that country
Figure 3. CDF of circuit risk over all countries, and against specific adversaries
Figure 4. Change in circuit risk, depending on whether the source/destination are within the adversarial country
Figure 5. Distribution of risk to different participants in Tor circuits.
Figure 6. Difference in risk between the guard and the source (US is the adversary)
Figure 7. Tabular presentation of MLAT country-level graph
Figure 8. Growth in Major Bilateral MLATs, 1961-2013, US, UK
Figure 9. MLATs signed by year, 1961-2013, US, UK
Figure 10. Change in Tor Users by country, 2011-2015, Greatest Increase
Figure 11. Change in Tor Users by country, 2011-2015, Greatest Decrease
Figure 12. Change in Proportion of Tor Clients by Country, 2011-2015
Figure 13: CAIDA’s autonomous system-level Internet graph
Figure 14. Graphic of GNI Funders, GNI Annual Report [135]
Figure 15. The MLAT.is web portal tool with database query and visualization
Figure 16. MLAT.is user statistics for top 10 countries as of August, 2017
Figure 17. Circuit risk if relays in the US are disallowed (g = U S)
Figure 18. Circuit risk if relays in Germany and MLAT allied countries are disallowed (g = DE)
Figure 19. Diameter of the Darknet: Minimum Spanning Tree Visualization
Figure 20. Diameter of the Darknet: Clustering visualization
Figure 21. Diameter of the Darknet: Thehub7gqe43miyc.onion visualization
Figure 22. Global MLAT cartel network visualization
Figure 23. Greg’s Cable Map [113] of Undersea Communications Cables
Figure 24: Dynamic Triggering of Lawful Intercept, Single Operator Common MF/CCTF Model, ETSI 102 677

List of Equations
Equation 1 country hostility metric
Equation 2 connection risk
Equation 2 circuit risk
Roadmap

We summarize the contributions of the various research projects and their artifacts resulting from this work: by contributions to knowledge, by research artifacts, by scientific or academic field, and by key findings. In the Overview, we provide an overview of all the related and subsequent research projects, and include a data dissemination and data management plan. In Section 1, we provide an introduction to work. In Section 2, we introduce Tor, how it works, attacks and our threat model. In Section 3, we present related work. In Section 4 we present our methodology and results. In section 5, we briefly review several subsequent research projects for which this works forms the basis, each of which have papers and tools in various stages of grant proposals, publication, submission, draft or implementation.

Figure 12. Change in Proportion of Tor Clients by Country, 2011-2015

“The US-Germany MLAT is the first United States MLAT to include special investigative techniques among permissible types of assistance. Specifically, Article 12 establishes that the Parties may use telecommunications surveillance... It is typical of our over 50 MLATs with countries around the world... It has several innovations, including provisions on special investigative techniques, such as telecommunications surveillance...” [84]
Statement of Mary Ellen Warlow, Director, Office Of International Affairs, Criminal Division, US Department of Justice: Hearing on Law Enforcement Treaties Before the Senate Committee On Foreign Relations, 109th Cong. 10, November 15, 2005

[We] recommen[d] that LEAs and governments...improve existing MLATs so that...they cover evolving IP-based communications services...includ[ing]...interception of electronic communications,...both communications content and communications data...” [131]


Figure 14. Graphic of GNI Funders, GNI Annual Report [135]

Overview

Network security attacks on the Internet have long compromised user security. Whether users realize it or not, some or all of their personal data and communications are vulnerable to access by third parties through many documented attacks, for example, through spyware. Governments may attack individual privacy, anonymity, and security for legitimate purposes, such as investigating criminals, or less legitimate purposes, such as silencing political opponents.

But those same governments have other, potentially more threatening, means of attack at their disposal to compromise network security. For example, through lawful intercept (LI), they can dispense altogether with spyware, and simply record all traffic passing through network elements located in their jurisdiction.

Users of anonymous communications networks like Tor seek to ensure privacy and anonymity. While Tor has proven valuable, which is why some governments censor its use, and despite years of research, today’s anonymous communication systems are still
vulnerable to passive deanonymization attacks orchestrated by powerful adversaries, such as Government Intelligence Agencies (GIAs). Existing work underestimates the threat posed by country-level attackers, since current models fail to account for legal analysis as well as technical analysis for collusion between allied countries via their legal jurisdiction over Internet resources. Further, existing proposals for improving anonymity systems are too abstract to be deployed. Many proposals call for incorporating “trust” into anonymous path creation algorithms, but none of these works define a quantifiable metric for ascertaining real-world trust.

This dissertation incorporates and summarizes research already published as well as research in various stages of grant proposals, papers submitted and in draft, a implemented database and a web portal tool incorporating database query and data visualization tools that has been used over 17,000 times in users in over 40,000 countries since implementation [144, 145, 146, 147, 148, 149, 150, 151, 152, 153].

It then presents the detail some of the research results from Jurisdictional Arbitrage [147]. The research investigates how to quantify trust in anonymous communication networks through metrics that are quantifiable, measurable, and readily available for all, or a significant subset of all, countries of the world. It identifies five LI threat or hostility factors which can indicate governmental risk to network communications through lawful intercept attacks, and includes gathering data pertaining to these factors into a database.

One of the five hostility factors identified is Mutual Legal Assistance Treaties (MLATs). These are a little-noticed legal tool that link countries of the world in an invisible web of interlocking law enforcement cartels. This dissertation investigates whether graph theory and tools may also be applied to analyze the MLAT network, and better understand its characteristics.

This research investigates whether graph theory and tools can help analyze traffic flow in these anonymous communications networks like Tor, identify jurisdiction-related vulnerabilities, and suggest solutions to improve them, in CircuitBlasTor [150]. For example, measures of centrality may reveal traffic concentration on a circuit path, where governments may optimally target lawful intercept attacks. Diffusing centrality could mitigate LI vulnerability, by suggesting alterations to the path selection algorithm.

In addition to being a possible trust factor for undermining user confidence in online privacy, anonymity, and security, MLATs provide their own attack vector for anonymous communications networks. MLATs enable legal pressure on CSPs worldwide, and so offer a significant vector to attempt to “break” Tor or its underlying encryption. Legal pressure on CSPs like Facebook, Microsoft and Google by foreign governments to locate surveillance-enabled servers in their countries, has led them to become deeply involved behind the scenes in the automation of remote global surveillance in the communications infrastructure through MLATs. Industry groups, like the International Chamber of Commerce (ICC), seek to incorporate technical standards for automated
surveillance LI into MLATs, through remote “dynamic triggering.” A multi-stakeholder group, the Global Network Initiative, funded principally by Microsoft and Google, announced on January 28, 2015 its public policy agenda, “Data Beyond Borders: Mutual Legal Assistance in the Internet Era,” [127] which sets forth a public policy agenda to further shape MLAT policy. This dissertation explores the concerns raised in this and other documents by empirically examining the pattern of MLATs through a graph theory analysis.

This dissertation includes the first database of international MLAT treaties [152] by identifying, documenting, and classifying thousands of MLATs. It further conducts the first analysis of MLATs to determine their actual impact on LI. We investigate two vectors to accomplish this task. First, a thorough review of MLAT case law in Legalizing Domestic Surveillance: The Role of Mutual Legal Assistance Treaties in Deanonymizing TorBrowser Technology [145]. Second, an investigation into how Google and Microsoft have become deeply involved in MLATs to further implementation of LI dynamic triggering, through international standards bodies such as the European Telecommunications Standards Institute (ETSI), for example in ETSI 102 877. Further, we make public policy recommendations for balancing interests in evolving MLATs in Dynamic Triggering: MLATs and Lawful Intercept in Communications Networks [148].

Returning to the five hostility factors as a whole, this dissertation develops unique metrics to quantify country hostility risk and circuit hostility risk. It uses the results of the evaluation, along with the results of our two-pronged graph theory analysis, to propose novel path selection algorithms and node placement to improve security, privacy and anonymity. Diffusing MLAT risk could also mitigate LI vulnerability, again, by suggesting alterations to the path selection algorithm.

MLAT.is [153] web portal is a new tool to better understand lawful intercept risk to network security, privacy and anonymity, analyze and map MLAT connections, and provide a means to understand their impact. This dissertation incorporates the data and research results from our lawful intercept trust metrics into MLAT.is, make it publicly available, and work to implement the results in the public and other Tor networks.

In MLAT: Internet Overlay [146], we build on the work in this dissertation to analyze the MLAT graph itself, demonstrating hidden cartels of collaborating countries which may use their legal powers of surveillance to attack user anonymity.

In CircuitBlasTor [150], we build on the work in this dissertation by then measuring millions of actual Tor circuits and analyze frequency of node selection, in comparison to Tor’s own node predictions. We demonstrate that Tor’s predictions err in some cases, creating a vulnerability for an attacker to control more traffic with fewer resources.

In Jurisdiction of the Darknet, Successful LEA De-Anonymization [149], we build on the work in this dissertation by analyzing criminal cases involving Tor Hidden Services. Hidden Services have legitimate uses, but have proven a haven for criminal activity over time. Vulnerabilities in Tor have enabled LEAs to deanonimize thousands of users in the last few years, and arrest and prosecute them. We review eight major cases and the technical and legal aspects of these cases.

Finally, in Diameter of the Darknet [151], we we build on the work in this dissertation by collecting a significant dataset of Tor Hidden Services. We crawl them and
collect their links. We then create a hyperlink map of the darknet and make key measurements like diameter, to produce a novel map of the darknet.

Keywords: Anonymous networks, internet communications, network security, privacy, anonymity, Mutual Legal Assistance Treaties (MLATs), surveillance, Tor, spyware, darknet, dark web.

1 Introduction to Jurisdictional Arbitrage

Network security attacks on the Internet have long compromised user security. Network researchers have long been skeptical about the security of communications on the Internet. Numerous papers assume the existence of an unseen, powerful adversary with the ability to eavesdrop, manipulate, or inject traffic on the Internet [15, 41, 104-110]. This mindset has been the driving motivation behind many advances in Internet security, such as anonymous communication systems [15, 25, 45, 56] and censorship resistant file sharing [10, 27, 70, 81].

Whether users realize it or not, some or all of their personal data and communications are vulnerable to accessed by third parties through many documented attacks, for example, through spyware [43]. Corporations may employ spyware to exploit user data for profit. However, harkening back to the original meaning of the term “spyware,” governments also employ this attack against privacy, anonymity, and security. They may use it for legitimate purposes, such as investigating criminals, or less legitimate purposes, such as silencing political opponents or launching active attacks against other countries. In this sense, spyware can be considered a weapon.

Today, we know that the threats to Internet traffic posed by large-scale, powerful adversaries are very real. China’s arsenal of Internet control and monitoring tools includes the Great Firewall [72], the hackers in Unit 61398 of the People’s Liberation Army [66], and crowdsourced workers in the “50 Cent Party” [85]. Iran has deployed sophisticated Deep Packet Inspection (DPI) devices, Trojan horses, and national firewalls to hinder access to foreign websites and spy on its citizens [46, 49]. These ongoing revelations about the scope of clandestine surveillance underscore that now, more than ever, systems that maintain the security, privacy, and anonymity of online communications are vital in the increasingly digitally-connected world.

Anecdotal evidence suggests that users are becoming more cognizant of threats posed by large-scale surveillance, and are taking steps to protect themselves. For example, the Tor anonymous communication system has been protecting users and slowly growing since 2002. In 2013 it grew by leaps and bounds: traffic flowing through the network increased by 75%, and the number of relay servers increased by 66% [2]. As of September 2015 usage is over 2 million daily tor client connections [2]. The news-media has speculated that more people are turning to systems like Tor to protect themselves online now that the threat of large-scale Internet surveillance has been laid bare [51].
But those same governments have other, potentially more threatening, means of attack at their disposal to compromise network security. For example, through lawful intercept (LI), they can dispense altogether with spyware, and simply record all traffic passing through network elements located in, or passing through, their jurisdiction.

Users of anonymous communications networks like Tor seek to ensure privacy and anonymity. However, while Tor has proven resilient, which is one reason some governments ban its use, and despite years of research, today’s anonymous communication systems are still vulnerable to threats posed by powerful adversaries. Although some solutions have been proposed to mitigate these attacks, these solutions fall short in multiple respects: First, existing work underestimates the resources available to, and therefore the danger posed by, country-level adversaries. Existing work assumes that the adversary may be able to monitor a small fraction of all ASes and IXPs, up to all of the infrastructure within a single country [17, 18, 41, 110]. Unfortunately, this model fails to account for the fact that allied countries can conduct cross-border surveillance. A great deal of evidence has come to light demonstrating that international surveillance is now widespread [26, 55, 60, 64, 78, 79, 84]. Second, solutions proposed in the literature for increasing the security of Tor circuits have shortcomings. One line of work focuses on selecting Tor relays in different ASes (to prevent a single AS from observing multiple hops) [17], but this fails to account for collaboration between adversarial ASes or countries.

Some have attempted to mitigate Tor’s vulnerability to attacks on privacy and anonymity, for example, by assigning trust to network elements such as nodes, or improving its path selection algorithms. We investigate whether it is possible to quantify trust in anonymous communication networks through metrics that are quantifiable, measurable and readily available for all, or a significant subset of all, countries of the world. We identify and analyze five hostility factors that meet these criteria that can indicate risk to network communications, of lawful intercept attack, for all countries. Once the basic pattern of hostility factors have been identified, this work can be enriched by substituting additional measures, or weighting these measures differently. This is a first attempt to model what such a quantification, if possible, would look like.

One of the five hostility factors I identify is MLATs. These are a little understood legal tool that link countries of the world in an invisible web of interlocking criminal law enforcement cartels. I demonstrate how MLATs directly impact lawful intercept.

In addition to comprising a possible trust factor for lawful intercept, MLATs provide their own attack vector for anonymous communications networks. We investigate MLATs, and compile a research database of thousands of treaties from dozens of sources. We construct a tool, MLAT.is, to analyze and map their connections, and provide a means to understand the impact of MLATs. We incorporate the data and research results into our lawful intercept trust metrics to help better understand lawful intercept risk to network security, privacy and anonymity. Diffusing MLAT risk could also mitigate LI vulnerability, again, by suggesting alterations to the path selection algorithm.
Returning to our five hostility factors as a whole, this dissertation develops unique metrics to quantify country hostility risk and circuit hostility risk. It uses the results of the evaluation to propose novel path selection and relay placement approaches to improve security, privacy, and anonymity.

2 Background

We first discuss background information on anonymous communication systems, beginning with Tor. Next, I discuss current research on attacks against Tor. Finally, I introduce our adversarial threat model.

2.1 The Public Tor Network

The Onion Router (“TOR”) Project was a research project initiated by the US Navy in the late 1990s. Syverson, Reed, and Goldschlag hold the 2001 Onion Routing patent on behalf of the US Navy [136]. It has given its name to a number of online communications networks, which use its schema as well as “borrowing” its title. For example, “The Tor Project” is a nonprofit organization incorporated in Massachusetts, USA in 2006 by Dingedine, Mathewson, Lewman et al. [137]. TOR, Syverson’s 1990s Navy project, gave rise to a number of onion routing networks. One of the best known of these is The Tor Project’s public “Tor” network. It should be noted that many other “TOR” networks are known to exist besides “Tor,” some of them private.

Figure 10. Change in Tor Users by country, 2011-2015, Greatest Increase

The public “Tor” network is an anonymous, online communications network. Tor is also the name given to a suite of free, open-source software that implements the TOR network protocols, backed by a global volunteer network of relays through which anonymized network traffic passes. The first alpha version of Tor was released in 2002 [13], with the first large-scale, public release following in 2004 [15]. As of September 2015, over 6,900 Tor relays handle over a 1.8 million users’ daily network traffic at an average rate of over 58 Gbits per second [2].
Figure 11. Change in Tor Users by country, 2011-2015, Greatest Decrease

Tor’s anonymity guarantees are based on the concept of TOR onion routing [63, 76]. When a user wishes to communicate anonymously with a destination, the Tor client software on the user’s computer constructs a circuit consisting of up to 10 relays. By default, three relays are used; in this case, the user’s packets are sent to the first relay, usually referred to as the guard. The client then sends the packets through the first relay to the second relay, and then repeats the process through the second relay by forwarding the packets to the exit relay. The exit relay sends the packets to the final destination. To communicate with the user, the destination may send packets to the exit relay, which can forward the packets to the user along the reverse path.

The confidentiality and integrity of packets are ensured by layers of encryption that are applied to the path itself at the user’s client, and subsequently removed connection-by-connection through each hop of the circuit. Conversely, the guard relay knows the source of data packets, but not their content or ultimate destination.

TOR encryption, applied by Tor, turns the path into a tunnel. It not only conceals the path, it further ensures that data packets only exist in the clear at the end of the circuit (i.e. from the exit relay to the destination.) Unless, of course, the user has applied their own, additional, non-Tor encryption to the packets themselves. Tor wraps the path at the source with three telescoped layers of encryption, creating the tunnel. Relays remove these, one layer at a time, to establish a path, the end of which does not know its origin. Once established at the destination, reply traffic to the source flows back on the same telescoping, tunnelled path.

To bootstrap the anonymity network, Tor provides nine public directory authorities that provide a census view of the 6,900 relays maintained by volunteers around the world. The directory contains the pertinent details about each relay, including its IP address, its measured uptime, whether the relay should be used as a guard, and whether the relay is willing to be an exit. Tor recommends that only relays with significant bandwidth and long uptimes be used as guards, since guard relays are of critical importance for
preserving anonymity (i.e. guards know the sources of circuits, since guards are the first relays on the path).

**Path Selection in Tor**

Selecting the three relays for an anonymous circuit is a key task in Tor [101, 102]. Each Tor client chooses one guard relay as the initial relay on all circuits. The guard is replaced periodically. This process is designed to prevent malicious relays from becoming guards. Intermediate relays and exit relays are chosen from the directory using a weighted random selection process where relays with longer uptimes and more available bandwidth receive higher weights. This process helps to spread traffic load over the relays, as well as balance anonymity versus network performance for clients. Tor clients change the relays in long-lived circuits every six minutes to hinder attempts to observe traffic.

Several papers have proposed alternate relay selection algorithms for Tor, and anonymous communications systems in general. This includes strategies that try 1) to avoid selecting relays in the same ASes (to improve resiliency against AS-level adversaries) [5, 17, 41, 110, 2] incorporate “trust” into the selection process [12, 40, 100], and 3) improve network performance by measuring the characteristics of relays [71, 77].

**Attacks Against Tor**

Tor, like other anonymous communication systems, is vulnerable to a variety of attacks that can be used to compromise a user’s anonymity. The earliest related work focused on developing entropy metrics for quantifying system anonymity guarantees [11, 69]. Later work takes a different approach by considering an attacker with bounded resources, and uses probabilistic analysis to quantify anonymity given the limits on the attacker (i.e. how many routers they can compromise at a single time) [75]. However, these works focus on abstract models of attacker capabilities, rather than focusing on actual network topologies or exploits that can be used to implement attacks in the real world.

Other work has examined more concrete threats to anonymous communication systems. Wright et al. propose the predecessor attack, in which malicious guard relays use statistical analysis to identify the sources of circuits [90]. Tor’s conservative guard selection process is designed to mitigate this threat. Numerous studies examine traffic correlation attacks (sometimes called the first-last correlation attack), where an attacker attempts to probabilistically deanonymize users by correlating traffic that enters and exits the anonymity network [53]. This attack has been shown to effectively deanonymize hidden services on Tor.

Other studies use latency measurements to correlate anonymous users observed at destination servers [28], or conduct low-resource attacks to alter network characteristics in ways that decrease anonymity [7, 8, 38], among others [20]. Several studies specifically examine deanonymization attacks against Tor hidden services, including statistical correlation attacks [57], and timing-attacks based on clock-skew [52]. And, others have proposed attacks related to anonymity and ASes, and routing generally [6, 42, 44, 47, 92].
**AS and IXP-Level Attacks**

Several studies have examined the impact of malicious Autonomous Systems (ASes) on anonymous communication systems [18, 110]. Feamster et al. were the first to observe that as a given anonymous circuit grows longer (i.e. the relays are more geographically diverse) the probability of having a single AS on both sides of the anonymous path increases [18]. This gives an AS-level adversary the ability to perform a traffic correlation attack. Measurements of the Tor network indicate that this pathological condition may impact 18% of circuits [18]. However, malicious ASes are not the only vantage point to conduct deanonymization attacks. Murdoch et al. further consider attacks launched by adversarial Internet Exchange Points (IXPs), i.e. the co-location facilities where one or more ASes peer with each other [54]. A malicious IXP is potentially more dangerous than a malicious AS because it can conduct correlation attacks on traffic from multiple ASs.

**2.2 Threat Model**

Before delving into the details of international law, we first define our threat model. In this work, we assume that the attacker is the government of a nation-state, or some agency acting on behalf of the government (e.g. NSA, GCHQ, Mossad, etc.). For simplicity, we simply refer to the adversary as a government intelligence agency or (GIAs). We assume that the attacker's goal is to deanonymize the communications of one or more Tor users.

Although there are many possible attacks against Tor, in this study we assume that the government's goal is to perform a (passive) traffic correlation attack. Thus, the GIA's modus operandi is to observe the traffic entering and exiting as many Tor relays as possible. An adversarial government can implement this attack through live traffic monitoring (e.g. a wiretap), or by appropriating historical traffic logs collected by a telecommunications infrastructure organization such as an ISP, AS, or IXP. Our overarching goal in this research is to understand, quantify, and counteract the threat to Tor posed by GIAs. These include the NSA in the United States, GCHQ in the United Kingdom, GRLS in Iceland, and FRA in Sweden.

We assume that the capabilities of the GIA are bounded by geographic and (potentially) legal limitations. In the worst case scenario, a government may be able to arbitrarily monitor, record, or alter all Internet traffic passing through ASes and IXPs within its legal jurisdiction. In other cases, a government's ability to monitor and record Internet traffic may be restricted by local laws and regulations. For example, in the US all wiretaps require a court order. Additionally, allied governments may collude to increase their monitoring and recording capabilities. We discuss the details of the international treaties that facilitate collusion in the next section.

To simplify our analysis, we conservatively assume that the adversarial government may target users regardless of their nationality or physical location. In reality, these two factors may may hinder the government's attack by preventing specific laws from being leveraged.
3 Related Work

This section provides an overview of related work in three fields. First, it provides the context of related work in computer science, in the field of network security attacks such as spyware and botnets. Second, it reviews the relevant work in network theory. Finally, it reviews work in the field of Law and Legal Studies, where research on MLATs is somewhat scarcer.

3.1 Computer Science: Network Security

The literature on threats and attacks to network security on the Internet is extensive. In addition to the Tor-specific attacks described in §2.1 above, for example, classic network security attacks include, for example, cross-site scripting [21], spyware [43] and botnets [22, 35, 36]. The commercial Finfisher surveillance software, used by many governments to spy on network traffic, utilizes Trojan spyware [39]. It employs a network security attack that enters through emails masquerading, for example, as an iTunes update [62].

3.2 Network Science: Communications and Lawful Intercept Cartels

Network theory is thought to have originated with Euler’s 1735 solution to the Seven Bridges of Königsberg problem [138]. With the emergence of computers in the late 1960s, commenced an era when it became feasible to analyze large datasets. Around this time, network theory took what would later be recognized, in combination with giant advances around after 1999 [93, 94, 95, 126], as the first steps towards the modern Network Science discipline that has emerged. Sonquist and Koenig, for example, were among the first to apply machine computing to questions of social networks, e.g. interlocking corporate directorships, 1975.

Network Science provides the means for performing this analysis, and, indeed, graph analysis has long been applied to the Internet [93, 94, 95, 96, 97, 126]. R and the Python libraries provided by NetworkX are two Network Science tools which can enable our analysis [133, 140].

According to Ulrik Brandes, to whom the NetworkX betweenness centrality algorithm is attributed, centrality can be defined as a vertex that “lies on considerable fractions of shortest paths connecting others” [132]. By this definition, target ASes achieve a score on the centrality list as surveillance targets by virtue of their position on shortest paths.
Many have graphed all ASes on the Internet, for example, CAIDA [134].

See CircuitBlasTor [150] for a more in-depth description of the application of Network Science to this research.

3.3 Law: Privacy, 4th Amendment, Search and Seizure, Intern’l Criminal Treaties

Mutual Legal Assistance Treaties, or MLATs, are little-noticed legal instruments that enable legal pressure on CSPs worldwide, and so offer a significant vector to attempt to “break” Tor or its underlying encryption. Legal pressure on CSPs like Facebook and Google by foreign governments to locate surveillance-enabled servers in their countries, has led CSPs to seek to incorporate into MLATs the automation of remote global surveillance. Their industry groups, like the International Chamber of Commerce (ICC), seek to incorporate technical standards for automated surveillance LI into MLATs, through remote “dynamic triggering.” A multi-stakeholder group, the Global Network Initiative, announced on January 28, 2015 its public policy agenda, “Data Beyond Borders: Mutual Legal Assistance in the Internet Era,” which sets forth a strategy to further shape MLAT policy. [127]

Surprisingly little is written about MLATs, which have succeeded in avoiding public notice, possibly due to their complexity and the lack of immediately obvious impact on the internet. However, there is substantial case law involving MLATs. In general, MLATs bear on a number of legal studies fields, including:

- Law and Technology
- International Law
- Cybercrime
- Surveillance Law - Lawful Intercept - Privacy Law
• Constitutional Law - Fourth Amendment- Search and Seizure

A complete survey of related work is incorporated into my legal publication surveying treaties and case law, including the September 30, 2014 and related Second Circuit rulings in US v. Getto, and the September 2013 First Circuit rulings in Boston College Trustees, Legalizing Domestic Surveillance: The Role of Mutual Legal Assistance Treaties in Deanonymizing TorBrowser Technology [145].

On related work in Law and Network Theory, Koenig first wrote for example, about interlocking Corporate Directorates and Network Theory in 1975 [65]. On International Law and Cybercrime, see Koenig [73, 82, 83]. See also section §4.2 for a more in-depth description of the application of Legal Studies to this research.

4 Research Methodology and Results

4.1 Jurisdictional Arbitrage: Quantifying Threats to Anonymous Communication Networks by GIAs

Now that we have defined the attacker and its goal, we examine the capabilities of adversarial governments in greater detail. Our first research objective is to develop a quantifiable understanding of GIA capabilities. We know that intelligence agencies are wiretapping the Internet [26, 60, 67, 79] and actively seeking to subvert Tor [68] currently, but we lack understanding of the capabilities of these GIAs. Furthermore, although existing work has identified the danger to Tor posed by AS and IXP-level surveillance [17, 18, 41, 110], we argue that these works underestimate the threat posed by GIAs. In particular, countries can collude to greatly expand their Internet surveillance capabilities. Currently we lack understanding about the web of international treaties that facilitate cross-border law enforcement and data sharing between GIAs.

First, this dissertation identifies empirical data that captures the technical and legal tools that are available to GIAs for conducting passive surveillance attacks against anonymous systems. As a starting point, I have identified and focus on five measures of LI risk, or trustworthiness, which I refer to for convenience as “hostility factors” that I describe in § 3.1.1. Second, I use our empirical data to develop measurement-driven metrics that quantify the risk posed by GIAs in each country to anonymous Internet traffic. We present proposed metrics based on our five trust factors in § 3.1.1. We ultimately seek to leverage these metrics to develop and evaluate new, more secure anonymous circuit creation algorithms. Third, and finally, we present preliminary results that motivate the efficacy of our approach.

Defining Hostility Factors and Collecting Hostility Factor Data

As a starting point for our research, I have identified and focus on five hostility factors that capture the laws, regulations, policies, and treaties that impact Internet surveillance around the world, and may indicate a government's willingness and ability to monitor and record Internet traffic:

• Passing Lawful Intercept (LI) laws or conducting surveillance using LI technology.
• Passing data retention laws.
• Conducting Internet censorship.
• Enforcing Internet censorship policies.
• Entering into treaties with other governments or Non-Governmental Organizations (NGOs) that facilitate cross-border law enforcement or information sharing.

We gather detailed data about the laws, regulations, policies, and treaties that impact Internet surveillance around the world. This data gives us a fine-grained view of the capabilities of individual governments with regards to attacking and deanonymizing Tor traffic.

We do not represent that these are perfect factors to measure hostility to anonymous network traffic. We have formulated metrics and algorithms so anyone can substitute their own. We seek, rather, to investigate whether it is possible to quantify and assign trust with any measure of reasonableness and if so, what might be such a framework?

Our criteria for selecting factors and measures are that they must be: Quantifiable, Available, Relevant and Comprehensive, covering all (or many) of the 249 ISO 3166 countries of the world. We call these QARC criteria. We reason that, if there are no QARC metrics, or they are not available worldwide for comparison, then we have already answered the question: it is not possible to assign trust based on worldwide lawful intercept threat.

We have conducted an extensive legal and literature search to identify laws, regulations, policies, and treaties that have bearing on these five hostility factors, which I have compiled into a database. However, given that there is no comprehensive repository for these kinds of legal data, I continue a comprehensive literature search for information on these hostility factors. (I have begun work on a sixth factor, dynamic triggering, which includes various national and international technical telecommunications and other internet standards, including European Telecommunications Standards Institute (ETSI) standards. ETSI TS 102 677 [98] details “dynamic triggering” for fully automated “Lawful Interception,” i.e., surveillance, incorporated into CSP technical design and infrastructure through MLATs. We have already begun to work with ETSI.) I plan to continue updating the database of knowledge regarding hostility factors as I uncover more relevant material over the course of this project, and as international laws evolve over time.

To keep our analysis tractable, I have thus far constrained our focus to the laws and policies enacted by the national governments of the 249 countries currently recognized by ISO 3166. This dissertation also considers bilateral and multilateral treaties between these countries, as well as laws pertaining to the EU. I do not consider laws or regulations issued below the national or federal level, e.g. state, provincial, or municipal laws. Unfortunately, in the real-world the jurisdictional boundaries of laws are often unclear. In many countries, the law of jurisdictional supersets, like the European Union (EU), or subsets, like state and even local law, can bear on network traffic. For example, in Germany, federal privacy laws govern network tracking. However, German state law may also be relevant, and each state has its own court authority similar in many ways to
the US Supreme Court. Likewise, as a member of the EU, EU privacy laws apply in Germany as well. We have taken a best-effort approach to resolving ambiguities in our database, however future legal precedents may necessitate alterations to our data.

I now describe the details of each of the hostility factors. When possible, I give concrete examples of laws or treaties in each category. I also provide a high-level overview of the database and data within each category.

**Internet Traffic Surveillance: “Lawful Intercept” (LI)**

The first hostility factor that I examine is government Internet traffic surveillance, otherwise known as “Lawful Intercept” (LI). LI refers to the collection of data from communication networks via means authorized by law. In essence, LI is official nomenclature for government-sponsored Internet Traffic Surveillance, or wiretapping. Users of anonymous communication systems (e.g. Tor) are threatened whenever their traffic passes through a country that implements Internet surveillance, since this monitoring can be used to implement passive deanonymization attacks. We have identified two features that indicate the state of surveillance in different countries.

The first is Lawful Intercept (LI) laws, or wiretap laws, that authorize government agencies to conduct surveillance on communication networks. Examples of LI laws include the Communications Assistance for Law Enforcement Act (CALEA) in the US, the Regulation of Investigatory Powers Act 2000 in the UK, and the Criminal Justice Mutual Assistance Act of 208 in Ireland. LI also encompasses surveillance tools, as well as mandates that hardware and software support surveillance capabilities (e.g. through APIs or backdoors for law enforcement). Users of anonymous communication networks (e.g. Tor) are threatened whenever their traffic passes through a country that implements LI. The Internet traffic monitoring enabled by LI is the essential component for implementing traffic correlation attacks.

In our database of hostility factors, I assign each country a binary attribute denoting whether that country has LI laws. Although it could be argued that some LI laws are stronger than others (e.g. more judicial oversight required when authorizing wiretaps), I make no attempt to quantify these subjective distinctions. We leave the creation of a ranked database of LI laws as future work.

However, I note that there are cases where governments install and use surveillance technology in the absence of LI laws explicitly preventing the GIA from doing so in that jurisdiction. Thus, anonymous communications may also be at risk when LI laws do not indicate hostility to anonymous communications. We have found that the best available source of information about actual LI implementations around the world is the BuggedPlanet Wiki [1]. This website compiles detailed information about LI laws, organizations known to leverage LI (e.g. Mossad), and known examples of the use of LI technologies.

Therefore, the second mechanism is the use of commercial or other software applications to conduct LI to spy on individuals. For example, Citizen Lab reported Finfisher had been observed in 36 countries around the world [50]. According to internal Finfisher documents dated 9/15/2014 (the latest data available as of 9/1/15), an additional
four governments are customers, for a total of 40 [86]. We view these two features as complementary, since the second case (actual software installations) allows us to classify many countries that do not have explicit LI laws. For our preliminary work, I have used data from BuggedPlanet to identify countries that have LI laws or use surveillance tools. Table 1 lists the top countries containing the most Tor relays, as well as whether these countries implement Internet surveillance (the LI column).

<table>
<thead>
<tr>
<th>Country</th>
<th># of Relays</th>
<th>% of Relays</th>
<th>% of Bandwidth</th>
<th>LI</th>
<th>DR</th>
<th>% of Takedown Requests</th>
<th># of MLATs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>1427</td>
<td>26.7%</td>
<td>24.1%</td>
<td>✓</td>
<td></td>
<td>10.7%</td>
<td>37</td>
</tr>
<tr>
<td>United States</td>
<td>1375</td>
<td>25.7%</td>
<td>19.5%</td>
<td>✓</td>
<td>✓</td>
<td>32.1%</td>
<td>87</td>
</tr>
<tr>
<td>France</td>
<td>339</td>
<td>6.3%</td>
<td>17.5%</td>
<td>✓</td>
<td>✓</td>
<td>8.5%</td>
<td>45</td>
</tr>
<tr>
<td>Russia</td>
<td>302</td>
<td>5.6%</td>
<td>1.4%</td>
<td>✓</td>
<td></td>
<td>0.1%</td>
<td>1</td>
</tr>
<tr>
<td>Netherlands</td>
<td>301</td>
<td>5.6%</td>
<td>10.0%</td>
<td>✓</td>
<td>✓</td>
<td>0.5%</td>
<td>43</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>214</td>
<td>4.0%</td>
<td>2.9%</td>
<td>✓</td>
<td>✓</td>
<td>8.6%</td>
<td>53</td>
</tr>
<tr>
<td>Sweden</td>
<td>173</td>
<td>3.2%</td>
<td>8.1%</td>
<td>✓</td>
<td>✓</td>
<td>0.2%</td>
<td>12</td>
</tr>
<tr>
<td>Canada</td>
<td>107</td>
<td>2.6%</td>
<td>1.4%</td>
<td>✓</td>
<td>✓</td>
<td>0.2%</td>
<td>16</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>96</td>
<td>1.8%</td>
<td>0.9%</td>
<td>✓</td>
<td>✓</td>
<td>0.1%</td>
<td>2</td>
</tr>
<tr>
<td>Switzerland</td>
<td>91</td>
<td>1.7%</td>
<td>3.1%</td>
<td>✓</td>
<td>✓</td>
<td>0.2%</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 1: Snapshot of our dataset for 10 countries with the most Tor relays

**Internet Traffic Surveillance: Data Retention Laws**

The second hostility factor under consideration is Data Retention (DR) laws. These laws compel ISPs and other communication service providers to store detailed records of network traffic and/or meta-data (i.e. IP addresses, telephone numbers, etc.) passing through their systems. The length of time data must be retained varies: The EU Data Retention Directive stipulated 6 month to 2 year storage [120], while Argentina's data retention law stipulated 10 years [121]. DR laws are a threat to anonymous communication networks because they enable adversarial governments to acquire long term, historical records of Internet traffic.

These long term records enable an attacker to perform large-scale traffic correlation attacks at their leisure. In contrast, LI only affords visibility into live traffic. However, LI and data retention laws are complementary, i.e. an adversarial government with both powers is more dangerous than a government with either power in isolation.

On April 8, 2014, the Court of Justice of the European Union (CJEU) declared the most far-reaching of these, the EU Data Retention Directive, to be invalid [9]. Nevertheless, as of September 2015, several active data retention laws continue in full force in EU countries, including, ironically, Ireland, the country which appealed to the CJEU. Argentina's data retention law was ruled unconstitutional in 2009. Brazil, and the US have all considered data retention laws, but they have yet been ratified. Earning it the moniker “Idiot Global Village” [19], Australia passed a sweeping data retention law in March, 2015 [16]. The Arbeitskreis Vorratsdatenspeicherung (German Working Group on Data Retention) [91] maintains and updates a log of worldwide Data Retention Laws, as does Open Rights Group [89], and EFF follows the issue as well [24].

However, revelations about the extent of NSA surveillance indicate that the there is de-facto data retention in the US. In particular, the NSA has permanent listening posts embedded in major ISPs (e.g. Verizon [122]) and IXPs (e.g. Global Crossing, which is now part of L3 [79]). The data syphoned from these listening posts are stored in massive data centers like the NSA's installation in Bluffdale, Utah [87]. Furthermore, the Stored
Communications Act (SCA) allows law enforcement in the US to subpoena records stored by communication network providers. Thus, even in the absence of mandatory data retention laws, the US government has enormous power to gather, store, and analyze records of Internet traffic. In our database, 12 countries are marked as having data retention laws. This includes many of the EU member states, plus others such as Denmark, Sweden and ironically, Ireland, as well as the US.

**Internet Censorship- Takedown Requests and User Monitoring**

The third hostility factor under consideration relates to the presence of Internet censorship. On the surface, censorship may not appear to be related to deanonymization: in fact, users often leverage tools like Tor to circumvent censorship [123]. However, government surveillance is a necessary precondition for implementing government censorship. Thus, I view censorship as a “canary in the coal mine” for surveillance. As discussed in the previous section, surveillance is definitely a threat to anonymous communication networks.

We leverage two sources of information to quantify Internet censorship by individual countries. First, data from major CSPs like Google and Microsoft in their “transparency reports.” We have collected transparency reports from Apple [31], Facebook [74], Google [32], Microsoft [33], Yahoo [34] and continue to research others [30]. These reports detail the content takedown and user data requests received by each company, broken down by country. This can be considered ex post facto censorship-attempting to close the barn door and get the horse back in, after it has already run out, so to speak. Takedowns refer to requests to remove content already published on the Web or have links delisted from search engines. Although some of the takedown requests are from copyright owners who want infringing content removed from the Web, others originate from government requests to censor information. For example, Google received requests from the governments of Turkey and Thailand asking for the removal of YouTube videos and blogs critical of their respective governments. These censorship requests reveal the willingness of countries to monitor and police content on the Web.

The user data requests in the transparency reports refer to attempts by law enforcement to acquire information about users of CSP services (e.g. user names, IP addresses, emails, Skype records, etc.). User data requests also indicate the willingness of governments to seek out information about Internet users. Table 1 lists the most and least active countries for takedown and user data requests.

**Internet Censorship- Direct Blocking**

The fourth hostility factor taken into account also relates to Internet censorship, in this case, pre-emptive rather than ex-post-facto censorship. We leverage Tor's anomaly-based censorship-detection system [88] as well as Tor’s Open Observatory of Network Interference (OONI) project [61] to determine which countries censor Internet traffic. Tor’s anomaly-based detection systems compares certain measures and ratios of users worldwide to determine censorship likelihood. OONI maintains a global, volunteer network of Internet measurement points that test for content blocking and traffic manipulation (e.g. DNS poisoning, injected TCP RST packets, HTTP header manipulation, etc.). In particular, OONI has issued detailed reports on censorship in
China, Ethiopia, Iran, Kazakhstan, Syria, the Philippines, and UAE. In our database, I record a binary flag for these seven countries, indicating that they have been flagged by OONI for Internet censorship.

**Internet Traffic Surveillance: Int’l Law Enforcement Treaties (MLATs)**

The fifth and final hostility factor is law enforcement and signal intelligence treaties between countries. These treaties facilitate cross-border collaboration between law enforcement agencies and data sharing between intelligence agencies.

Politicians have openly acknowledged that MLATs are being used to facilitate cross-border surveillance [84]. In particular, the NSA has been able to conduct surveillance in Germany, Brazil, and Australia [116] as well as Sweden [99] thanks to bilateral treaties. In fact, the Director of International Affairs of the US Department of Justice testified before the Senate Foreign Relations Committee in 2005 that the US-Germany MLAT was “the first United States MLAT to include special investigative techniques among permissible types of assistance. Specifically, Article 12 establishes that the Parties may use telecommunications surveillance. . .” [84]. She noted that it is.”...typical of our over 50 MLATs with countries around the world...It has several innovations, including provisions on special investigative techniques, such as telecommunications surveillance...” [84]. These revelations confirm that international treaties are being used for massive Internet surveillance, and that in some cases, lawmakers are specifically designing these agreements to facilitate this surveillance.

Some of these treaties are bilateral, i.e. formed between two countries. For example, the US has entered into Mutual Legal Assistance Treaties (MLATs) with 69 countries [124]. The Organization for Economic Co-operation and Development (OECD) has promulgated bilateral MLATs between most of its 34 member countries [125].

Multilateral law enforcement and signal intelligence treaties include The USA-UK Agreement, an intelligence sharing covenant between the “Five Eyes,” i.e. US, the UK, New Zealand, Canada, and Australia [116]. The Wassenaar Agreement is an arms and dual-use technology export control treaty between 42 countries that regulates the transfer of LI technologies around the world [115]. Finally, perhaps the most far-reaching treaty in this area is the Budapest Convention on Cybercrime. This treaty stipulates that ratifying countries must standardize their laws on cybercrime (e.g. copyright infringement, fraud, and network security) and cooperate with international investigations into cybercrimes. As of September 2015, 47 countries, including the US and most members of the EU, have ratified this treaty, and another 7 signatories are pending [103].

For this work, I create a graph of country-country law enforcement and signal intelligence treaties. In this graph, each node is a country, while each edge is an MLAT or other treaty between two countries. Our preliminary MLAT graph contains 249 countries connected by 1,192 treaty-edges. Figure 7 provides a tabular presentation of the MLAT country-level graph, while Figure 6 shows country-node, treaty-edges in a map graphic.

These MLATs and multilateral treaties enable adversarial governments to collude in deanonymization attacks. Tor frequently builds circuits that include relays in different countries. If the guard and exit relays are both within a single country, then that
government can conduct a traffic correlation attack by monitoring ASs and IXPs in its jurisdiction. However, in a multi-country, multi-government scenario, governments with jurisdiction over the multiple relays of a Tor circuit need to collude to mount an effective attack.

In the database, I have built a table of all known country-country law enforcement and signal intelligence treaties. For each pair of countries, I store a binary value denoting whether there is a treaty between those two countries. Theoretically there are $2^{49^2}$ possible treaty pairs, or over 62,000/2. However, in reality, 69 of the ISO 3166 recognized states are dependent territories, leaving $180^2$ or just over 32,000/2 possibilities. In total, we are aware of 1,192 such relationships, out of all possible country-country pairs.

**Challenges**

MLAT research is comparatively new among privacy experts and legal scholars, especially on a comprehensive, worldwide basis. In our efforts to collaborate with organizations like EFF, ACLU and Privacy International, these groups expressed a lack of comprehensive view or expertise in MLATs. Similarly, while organizations exist with some expertise in data retention laws for certain geographic areas, none of the most respected organizations had the resources or focus to develop expertise in the comprehensive global view required by this research.

**4.2 MLAT Database: Documenting and Classifying worldwide MLATs**

This task involved researching and documenting the thousands of worldwide MLATs. The result is embodied in the MLAT.is tool, now publicly available, described in 4.6, and has been access over 17,000 times since publication.
First, I define countries. Second, I identify original sources of treaties and treaty data for thousands of treaties. Third, I classify them according to various factors affecting their risk. This database is publicly available through the MLAT.is tool. It has been used in several subsequent works [144-151, 153].

**Country Definition Data**

We start with the International Organization for Standardization (ISO) 3166 standard to identify autonomous governments of the world. ISO, founded 1947, counts 164 member countries. As of the latest release, ISO 3166 recognizes 249 “countries.” However, ISO 3166 includes dependent territories with governments in various stages of control by other governments, which I refer to as “legal parents.” We therefore further identify 69 dependent territories using a number of sources. Classifying these along a spectrum of self-governance, I focus on the remaining territories, which I describe as “independent territories.” However, I maintain legal status relationship records of all dependent territories, including 48 with ISO 3166 country codes and 21 without country codes. We correlate 65 of the 69 dependent territories with their legal country “parent,” and maintain a list for each parent country of its “children” countries and therefore, treaty relationships. The remaining four territories include Hong Kong and Macao, associated informally with China, and Cook Islands and Niue, associated informally with New Zealand.
MLA Treaty Source Data

Treaty data itself comes from a large number of original sources, including governments and their agencies, such as the U.S. State Department, and global organizations, such as the Council of Europe (“COE”). Treaty source forms include lists of signatories by signatory nations as well as the treaties themselves.

MLA Treaty Metadata and other Data

Other data relevant to treaties includes authoritative country records of their own MLATs. Many countries keep and make them publicly available metadata summaries of their MLATs, including how many and with whom, when drafted and signed, and other information. Comparing countries’ own treaty metadata against the treaties themselves often reveals missing treaties or data, or conflicting information.

Treaties often include addendum documents and appendices which are incorporated by reference into the treaty. As a crucial example, the US-Russia MLAT not only did NOT exclude political offenses, it appends a nice note from the US Embassy agreeing it had specifically omitted such during negotiations.

Furthermore, I locate, track and store transcripts of legislative hearings which document the original intent and representations of MLAT champions. Since US case law regarding treaty interpretation often revolves around original intent of drafters, transcripts of legislative hearings often shed more light on the meaning and intent of treaty provisions, than the treaty texts themselves.

Research and Data Challenges

In researching, building and implementing the database, a number of challenges present themselves. For example, finding treaties, authenticating treaties and document versions, verifying and authenticating letters, notes and other official appendices to treaties, identifying and tracking the many dates associated with treaties, all present their own issues. Treaties by nature are signed between two or more parties. Each one may store their own copy, possibly in their own (non-English) language. How to authenticate each translation and version? Most treaties require a copy to be “deposited,” once fully approved, ratified and executed, in a “repository” in order to enter into force. This other “definitive” copy of the treaty may be made available online, possibly in yet another language. The records of the various repositories may be more or less up-to-date, providing contradictory information.
Not all treaties are readily available. Despite the US requirement for the State Department to publish all treaties, some remain stubbornly difficult to locate. Yet, a record of them may appear in reliable, definitive lists of treaties in force. Thus, MLAT research presents an ongoing series of challenges.

Even if treaties are identified, authenticated and located, their status is often difficult to ascertain and may change. Treaties have many phases through which they may pass on their way to being “In Force,” which can span over a decade or more. Some phases include: drafting, signing, acceptance/approval/ ratification/ accession (which can all be different, or the same), deposit of instrument, and entry into force (EIF). Indeed, some MLATs were drafted years ago but have never entered into force. For many database purposes, a treaty relationship between two countries either exists or it doesn’t, i.e. there is either a treaty in force (TIF), or there is not. Thus the existence of a treaty document NOT in force may lend confusion and lack of clarity.

Once treaties are located, verified and authenticated, their reading and interpretation can present challenges. Many provisions are brief and cryptic. Only upon location of legislative hearings do the real implications emerge. For example, the provisions that ultimately purport to prevent FOIAs appear, upon initial reading, impossible to decipher. Upon reading the full legislative hearing, for example, in the US, before the Senate Foreign Relations Committee, is it possible to understand that a certain provision was explicitly written to prevent FOIAs.

**MLAT Data Classification**

There are multiple kinds of MLATs, and their classification may be viewed along multiple vectors. Initially, I select several vectors for measurement and discussion, including: 1) Signatory Scope; 2) Category of Law; 3) Law Enforcement Goal; 4) Treaty Focus, 5) Criminal Threat Model. 6) Phase; and 7) Status.

Within those seven basic classification vectors, I identify multiple sub-options. In the first vector, signatory scope may be viewed as: multilateral, either global or regional; bilateral; or monolateral. Bilateral MLATs enshrine agreements between two countries, for example, the UK-Saudi Arabia MLAT.; multilateral MLATs, between more than two countries. Monolaterals may be considered as “one-way” laws enacted by
some countries enabling them to provide foreign MLA, in the absence of actual bilateral agreements with other countries. Regional multilateral MLATs may involve only a handful of countries, while global multilateral MLATs may involve almost every country. A country may enter into both multilateral MLATs with another country, and a bilateral treaty with that country, or even multiple bilaterals with a single other country, creating legal overlap. Some organizations administering multilateral MLATs include: the Organization for Economic Cooperation and Development (“OECD”), the Association of Southeast Asian Nations (“ASEAN”), the Organization of American States (“OAS”), the Economic Community Of West African States (“ECOWAS”), the European Union (“EU”) and the EU’s European Council (EU Council), the Council of Europe (“COE”), and United Nations (“UN”).

Naturally, not all MLATs are created equal. Some multilaterals are not directly enforceable in the member states. Instead, signatories are required to adopt laws or procedures which have the stated effect and are enforceable. For example, in the 1988 Vienna Narcotics convention, signatories are required to implement laws or procedures which will enable the signatories' own authorities to "identify, seize and freeze" proceeds or property derived from illegal drug activities. For this reason, bilateral treaties may be considered significantly stronger than many multilateral treaties.

Second, a treaty’s category of law may be either civil or criminal, as some MLATs deal with civil law and others, with criminal law. We concern ourselves principally with criminal MLATs or MLAT provisions.

Third, the law enforcement goal of treaties may include extradition (or “rendition”), or non-extradition mutual legal assistance (MLA). Extradition or Rendition treaties are often viewed as separate from MLATs, but they are closely related. We exclude extradition treaties from our analysis at this time, although I identify them as an area for future analysis.

In the fourth category, treaty focus, a treaty may either be simply a general MLAT, or it may have MLA provisions incidental to other treaty focus. Examples include tax agreements, customs agreements, Hague convention on child abduction, and Schengen Acquis Treaty- EU (abolition of internal border controls and a common visa policy).

Motivation for a number of MLATs has been linked to campaigns against specific crimes. These MLATs may be narrow, and may overlap with other MLATs. Criminal threat model, the fifth MLAT classification vector, may include one or more specific criminal threats. These may include: Narcotics, Anti-Bribery, Human Trafficking, Skyjacking, organized Crime, Terrorism, Weapons Trafficking, including nuclear weapons, or tax evasion incident to banking secrecy laws.

MLATs typically may take five to ten years between these four principal phases. For this reason, attempts to identify treaties by date can result in greater confusion. We
classify MLATs by a sixth vector, phase, to clarify whether and where it may be in force. Phases include: drafting, signing, acceptance/approval/ ratification/accession- can all be different, or the same, deposit of instrument, entry into force (EIF).

Only in the last of these phases is the MLAT actually in effect, a status often referred to as “Treaty in Force” (“TIF”). Likewise, treaties may be superseded by other treaties, a status which is not apparent solely from reading the treaty. For this and other reasons, such as expiration or a country ceasing to exist, a treaty may have a status or Not In Force (NIF). The seventh classification category, may include either Treaty in Force (TIF) or Not In Force (NIF). Another issue with MLATs is finding definitive sources, and agreeing on a definitive citation. As they are by nature signed by at least two parties, each may claim to hold the definitive reference. And, each treaty has multiple versions on its long journey from drafting to EIF.

4.3 Jurisdictional Arbitrage: Defining Metrics for Quantifying Risk from GIAs

In this section, I create a metric incorporating these hostility factors that can be used to predict the danger to a given Tor circuit from GIAs. In essence, the metric assigns a Tor circuit score based on 1) which GIA is attempting the deanonymize the circuit, 2) the countries that the source, relays, and destination are located in, and 3) the MLATs between those countries and the GIA’s country. The advantage of using country-level information to measure risk is that it is conceptually straightforward, and does not require knowledge of Internet topology. Our metric is divided into two components, which I discuss separately. First, I introduce the country hostility metric, which captures the hostility of a particular country towards Internet traffic. Country hostility incorporates four of our hostility factors: traffic interception, data retention, takedown requests, and censorship. Second, I introduce the circuit risk metric, which incorporates country hostility scores and MLAT relationships, and standards to predict the danger to a particular Tor circuit.

Taking into consideration country hostility indicator sources, I create a LI Threat or Legal Hostility rating algorithm and matrix. This framework summarizes all country hostility ratings in relation to all other countries, including a hostility rating to the privacy of anonymous network traffic of its own citizens, by applying an algorithm for determining each country’s hostility rating in relation to each other country’s network traffic. Each of the 249 countries identified by ISO 3166 has a different level of LI threat, or “hostility” to the anonymity of network traffic from every other country, which results in a 2,392 matrix of hostility ratings.

Country Hostility Metric

Our first goal is to distill the hostility factors introduced in § 4.1.1, into a quantifiable metric that captures the likelihood of Internet traffic being recorded or monitored as it passes through a given country. The higher this metric, the higher the likelihood of surveillance in that country, and therefore greater danger to anonymous
network communications. To develop our country hostility metric, I focus on four of the hostility factors from § 4.1.1: traffic interception, data retention, takedown requests, and censorship. Each one of these factors captures country-specific laws and practices. For a country $c$, let $I_{fc}(f = LI, DR, TD, CE)$ denote these four factors.

We define $0 \leq I_{fc} \leq 1$. For traffic interception, data retention, and censorship, this mapping is natural, since each of these three factors is binary. For example, a country either implements data retention or it does not. However, as shown in Table 1, takedown requests are positive integers. To convert takedown requests into a value between 0 and 1, I sort the countries from most to least takedowns, and assign them a score from 1 to 0 based on rank order, to ensure evenly distributed scores between 0 and 1.

We now introduce the country hostility metric $H_c$:

$$H_c = \sum \alpha_f I_{fc}$$

(1)

Simply put, the hostility $H_c$ of country $c$ is the sum of the four individual hostility factors $I_{fc}$. $\alpha$ is a weight associated with each factor $f$, defined such that $\sum_{\forall f} \alpha_f = 1$. Thus, the vector $\alpha$ serves two purposes: it ensures that $0 \leq H_c \leq 1$, and it allows the relative importance of each to be independently adjusted.

**Raw hostility score of country $c$**

Let $H_c$ = the hostility score of country $c$. Then let $RH_c$ = the raw hostility score of country $c$. $R$ is the sum of the four factors: $RH_c = \sum (LI, DR, TD, CE))$. For example, if ($LI= 1$, $DR= 1$, $TD = .72$, $CE= 0$), then $RH_c = \sum (1 , 1, .72, 0)) = 2.72$

**Scaling Between 0 and 1**

Scaling between 0 and 1 facilitates comparisons of the 249 countries of the world. If $\alpha$ is the total number of hostility factors, then, to ensure scores are evenly distributed between 0 and 1, I could simply divide each of the scores by $\alpha$. This provides for flexibility to incorporate additional hostility factors as they become available or change over time. Currently $\alpha = 4$, so

$$H_c = 2.72/4 = .68$$

Thus,

$$H_c = (LI + DR + TD + CE))/\alpha ; \text{ or}$$

$$H_c = \sum (LI , DR, TD, CE))/\alpha$$

**Weighted factors**

But what if each factor is not equally significant? The above formula can be expressed as

$$H_c = \sum (LI/\alpha, DR/\alpha, TD/\alpha, CE/\alpha)$$

where $\alpha$ serves the dual purpose to scale and to weight. We can then assign different values to $\alpha$ for weights. Our formula then becomes,

$$H_c = \sum (w_iLI, w_jDR, w_kTD, w_lCE)$$
where $0 \leq \sum (w_{i,a}) \leq 1$

In other words, the weights of each of the four factors must themselves add up to no more than one, to maintain our even distribution of scores between 0 and 1. We start by weighting all the factors the same, where $w_{i,a} = [1/\alpha, 1/\alpha, 1/\alpha, 1/\alpha]$, for example, $[0.25, 0.25, 0.25, 0.25]$ which keeps our scale between 0 and 1:

$$0 \leq \sum (1/w_{i,a}) = \leq 1$$

We can easily change the weights, if appropriate, for example,

$$w_{i,a} = [0.20, 0.30, 0.10, 0.40]$$

So that

$$H_c = \sum (0.20*LI, 0.30*DR, 0.10*TD, 0.40*CE)$$

In the first example, using the original hostility factors of (LI= 1, DR= 1, TD = .72, CE= 0):

$$H_c = \sum (0.20*1, 0.30*1, 0.10*.72, 0.40*0)) = .572$$

rather than .68 above

**Connection Risk Metric**

Now that a metric quantifies the hostility of a particular country, I move on to the second goal: quantifying the risk to an entire Tor circuit. This scenario is more complex because it involves a traffic source, several Tor relays, and a traffic destination that may be in different countries. To build our metric, I incorporate the fifth hostility factor discussed in § 4.1: MLATs between countries.

Suppose that a GIA in country $g$ wants to deanonymize a Tor circuit $C$ with any number of participants located in countries $c_1...c_n$. Let $C = [c_1, c_2...c_n]$. In this formulation, the source of the traffic is $c_i$, the destination is in $c_n$, and the traffic passes through countries $c_j$ through $c_{n-1}$ containing Tor relays. Intuitively, a GIA’s ability to conduct a passive attack against the circuit $C$ depends on two factors: 1) whether each participant $c_i$ is being monitored, corresponding to the hostility of country $H_{ci}$, and 2) whether the GIA $g$ has access to surveillance data in $c_i$, corresponding to the MLATs between the countries.

The GIA may have access because $g = c_i$, or because there is an MLAT between $g$ and $c_i$.

Based on these observations, I define Adversary Adjusted Risk (ARR) as $A_{g,c} = M_{g,c} * H_c$. $M_{g,c}$ encodes the MLAT relationship between the GIA country $g$ and country $c$. The purpose of $M_{g,c}$ is to adjust the hostility score $H_c$ of each participant country based on its relationship to the GIA country $g$. $M_{g,c}$ is defined over the MLAT graph $G = (V,E)$, where each $c \in V$ is a country, and each $e_{ij} \in E$ indicates that an MLAT exists between countries $i$ and $j$. We define $M_{g,c}$ as:

$$M_{g,c} = \begin{cases} 
1 & \text{if } g = c \text{ on } G \\
\frac{1}{Dg, c + 1} & \text{if a path exists from } g \text{ to } c \text{ on } G \\
0 & \text{if no path exists from } g \text{ to } c \text{ on } G
\end{cases}$$

$$\begin{align*}
S = A_{g,c} & = M_{g,c} * H_c \\
& = \frac{1}{Dg, c + 1} * H_c
\end{align*}$$

33
where $D_{g,c}$ is the shortest path length between $g$ and $c$ on the MLAT graph $G$. In the simple case where $g = c$, the path length is zero, thus $M_{g,c} = 1$ and risk is entirely captured by $H_c$. In cases where the GIA must collude with other countries (i.e. $g \neq c$), risk decreases proportionally to the path distance $D_{g,c}$ between $g$ and $c$. This corresponds to the intuition that it easier to acquire surveillance data from a direct treaty partner ($D_{g,c} = 1$) than an indirect partner ($D_{g,c} > 1$). Finally, if there is no way for $g$ and $c$ to cooperate then $M_{g,c} = 0$, which nullifies any risk posed by $H_c$.

Before quantifying the risk to an entire circuit, I first focus on the risk of surveillance to a connection between two participants in a Tor circuit. A connection is a pair of participants that communicate directly, e.g. the source and the guard relay, or the exit relay and the destination. Participants that are connected form an equivalence class for the purposes of surveillance. For example, an attacker that can observe the traffic leaving the source has the same power as an attacker that can observe the traffic arriving at the guard relay. Thus, for an adversary $g$ attacking a circuit $C$, I define connection risk as

$$X_i(g, C) = \max(M_{g,c}, H_c, M_{g,c+1} H_{ci+1})$$  \hspace{1cm} (2)

Thus, I assume a worst-case scenario where the risk to each connection is the maximum of the risk of surveillance at the two endpoints. The connection risk is the max of the hostility factor of source country and the guard country, times the MLAT factors that indicates whether an MLAT exists between the two countries.

**Circuit Risk Metric**

For example, where the five nodes are US -> France -> China -> Netherlands -> India, representing the source, guard relay, middle relay, exit relay, and destination, the relevant factor and weights include:

- Country Hostility: 0.75, 0.74, 0.66, 0.66, 0.48
- MLAT Degree: 1, 1, 0, 0
- connection weight: $\beta = [0.5, 0.0, 0.0, 0.5]

The circuit risk is equal to the sum of the four connection risks:

$$C(g,c) = X_1 + X_2 + X_3 + X_4$$

We adjust this circuit risk to account for different weight of the different positions:

$$C(g,c) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

We now define our circuit risk metric $\Psi(g,C)$ in terms of connection risk:

$$\Psi(g,C) = \sum_{i=1}^{n-1} (\beta_i \max(M_{g,ci}, H_c, M_{g,ci+1} H_{ci+1}))$$  \hspace{1cm} (3)
The risk $\Psi(g,C)$ posed by a GIA in country $g$ to a circuit $C$ is the sum of the connection risks $X$, adjusted by some weight $\beta$. Similar to the $\alpha$ vector in the country hostility metric, the $\beta$ weight vector is defined as $\sum_{i=1}^{n-1} \beta_i = 1$, and serves to ensure that $0 \leq \Psi(g,C) \leq 1$. $\beta$ also allows connections in different positions to be given different weights, especially the source/guard and exit/destination connections.

4.4 Jurisdictional Arbitrage: Applying Risk Metrics to Countries and Circuits

4.4.1 Evaluating Country Hostility

Evaluating Country Hostility. Now that the country hostility and circuit risk metrics are defined, we can evaluate the danger posed by GIAs in different countries. We begin by examining the hostility of individual countries, $H_c$. For simplicities sake, we set $\alpha = [0.25, 0.25, 0.25, 0.25]$, i.e. four hostility factors receive equal weight.

![Figure 1: CDF of country hostility scores.](image)

Figure 1 shows the CDF of nonzero hostility scores $H_c$ for ISO 3166-1 countries. The distribution shows a distinct stepped pattern, which corresponds to the four hostility factors in the equation. 47% of countries receive a hostility score of 0.0 because they do not display any of the four hostility factors. Examples of countries where $H_c = 0.0$ include many African nations and small, island nations. The noticeable inflection points at 0.25 and 0.50 hostility correspond to countries that exhibit exactly one or two (respectively) binary hostility factors (i.e. traffic interception, data retention, censorship). The escalating regions in between $0.0 - 0.25$, $0.25 - 0.5$, and $0.5 - 0.75$ are occupied by countries that sent takedown requests. Table 2 lists the 21 most hostile countries according to our metric.
Table 2: Top 21 most hostile countries.

<table>
<thead>
<tr>
<th>Country</th>
<th>$H_c$</th>
<th>Country</th>
<th>$H_c$</th>
<th>Country</th>
<th>$H_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>0.75</td>
<td>Netherlands</td>
<td>0.66</td>
<td>Estonia</td>
<td>0.52</td>
</tr>
<tr>
<td>UK</td>
<td>0.74</td>
<td>China</td>
<td>0.66</td>
<td>UAE</td>
<td>0.51</td>
</tr>
<tr>
<td>France</td>
<td>0.74</td>
<td>Hungary</td>
<td>0.65</td>
<td>Latvia</td>
<td>0.5</td>
</tr>
<tr>
<td>Italy</td>
<td>0.72</td>
<td>Denmark</td>
<td>0.63</td>
<td>Kazakhstan</td>
<td>0.5</td>
</tr>
<tr>
<td>Spain</td>
<td>0.71</td>
<td>Ireland</td>
<td>0.62</td>
<td>Ethiopia</td>
<td>0.5</td>
</tr>
<tr>
<td>Poland</td>
<td>0.69</td>
<td>Austria</td>
<td>0.59</td>
<td>Germany</td>
<td>0.5</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.68</td>
<td>Lithuania</td>
<td>0.54</td>
<td>India</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Riskiness of Existing Tor Relays

Our country hostility metric allows us to address the question, *are Tor relays located in countries with high risk (low trust) scores?* Indeed, as shown in Figure 2, the countries that host the most relays tend to also have $H_c \geq 0.5$. Figure 2(a) is a scatter plot showing the overall number of Tor relays in a given country versus its hostility score. Figures 2(b) and 2(c) focus on just guard and exit relays, respectively. Although only 20 countries have $H_c \geq 0.5$, 80% of relays are located in these countries (79% of guards, and 63% of exits).

![Figure 2](image)

*Figure 2. Hostility of countries versus the number of Tor relays in that country.*

![Figure 3](image)

*Figure 3. CDF of circuit risk over all countries, and against specific adversaries.*

4.4.2 Evaluating Circuit Risk

Next, we examine the danger of passive surveillance attacks against Tor circuits using our circuit risk metric $\mathcal{R}(g, C)$, defined by Equation 3. We leverage the country hostility scores $H_c$ presented in § 4.4.1, and $G$, the graph of MLAT relationships between countries introduced in §4.2. Tor defaults to selecting three relays per circuit, so I fix $n = 5$ (i.e. the source, three relays, and the destination) and set $\beta = [0.5, 0.0, 0.0, 0.0, 0.5]$. This weighting reflects the fact that observing the start and end of the circuit (i.e. the connection between the source and
the guard, and between the exit and the destination) is more important for the attacker’s success than observing the connection in the middle [41, 53, 57].

We leverage the Tor Path Simulator (TorPS) [111] to generate realistic Tor circuits for our evaluation. We examine the riskiness of Tor circuits by plotting the distribution of circuit risk for one million circuits generated by TorPS during a month in 2014. The country-specific lines in Figure 3 correspond to the risk distribution of the one million circuits versus particular adversaries. The US has the most MLATs (75), so it is not surprising that circuits are most vulnerable to monitoring by the US. A more interesting comparison is between Germany and China: although Germany has lower hostility than China ($H_{DE} = 0.5$ vs. $H_{CN} = 0.66$) and fewer MLATs (37 vs. 56).

Next, we examine the danger of passive surveillance attacks against Tor circuits using our circuit risk metric $\mathcal{H}(g,C)$, defined by Equation 3. We leverage the country hostility scores $H_c$ presented in § 5.1, and $G$, the graph of MLAT relationships between countries introduced in § 3.2.5.

![Figure 3: CDF of circuit risk against all adversaries, and against specific adversaries.](image)

**Key Findings and Observations**

- 71% of circuits in our dataset are at high risk of surveillance by the US.
- Although Germany has lower hostility than China ($H_{DE} = 0.5$ vs. $H_{CN} = 0.66$) and fewer MLATs (37 vs. 56), Figure 3 indicates that circuits are more at risk to passive attacks from the BND than the MSS. This result is due to the fact that, as of the experiment date, Germany hosts the most Tor relays of any country (see Table 1), whereas China hosts zero relays.
- When destinations are random, the risk of the last connection is dominated by the risk at the exit.
Figure 4: Change in circuit risk, depending on whether the source/destination are within the adversarial country

- There are many cases when the source and destination countries are less risky than the countries containing the guard and exit relays
- Thus, the risk at the guard must be less than or equal to risk at the source.
- 95th percentile risk at the source is only 0.32, while median risk at the guard is between 0.33 and 0.37. This illustrates that for most countries, risk at the source is significantly lower than risk at the guard across four powerful adversaries.
- We have confirmed that Tor’s relay selection algorithm sometimes chooses guard and exit relays that increase the risk of surveillance.

Considering these results, the following key questions arise:

- How frequently does the Tor client build pathological circuits, e.g. where guard risk is higher than source risk?
- How many countries are impacted by pathological circuit construction?
- What fraction of Tor’s users are impacted by pathological circuit construction?

Our observations and conclusions:

- We observe that the difference is > 0 (i.e. the guard country is riskier) in 95% of circuits. In 13% of circuits the difference is 0.75, meaning that source risk is zero, versus guard country risk of 0.75 (which occurs when the guard is in the US).
- When the source is in the US or the UK, source risk is always greater than or equal to guard risk. This is to be expected: the US and the UK are the two most hostile countries in our model (see Table 2). Conversely, when the source is in Japan or Taiwan, 83% and 100% of circuits (respectively) have riskier guards than sources.
- Our experiments reveal that in 201 countries (out of 249), 50% of circuits have guards in riskier countries than the source. Tor Metrics [2] provides estimates for the number of unique Tor users in each country each month. According to their user statistics for January 2014, 61% of Tor users reside in countries affected by this pathology. In other words, for the majority of Tor users, the Tor relay selection algorithm is likely to choose a guard node in a country with higher risk of surveillance than the user’s home country (assuming the most powerful adversary, the US).


**Circuit Dataset**

The Tor circuits used in our evaluation are generated by the Tor Path Simulator (TorPS) [111]. TorPS was developed to enable researchers to produce realistic Tor circuits for use in studying the Tor network. TorPS is built around three components: a Tor network model, a user model, and a Tor client model. The Tor network model captures the state of the Tor network (i.e. relays and their characteristics) at given points in time by taking as input server descriptors and consensus files from Tor Metrics [2]. The user model describes the activities of simulated users (i.e. the traffic they send, to whom, on what ports, and at what time), which drive the creation of Tor circuits. Finally, the TorPS client model implements Tor’s relay selection algorithm (as of Tor version 0.2.3.25). Taken together, these three components allow researchers to produce an unlimited number of realistic Tor circuits.

In this work, we use the following parameters with TorPS. We use server descriptor and consensus files for December 2013. In keeping with [111], relays that are not marked as FAST and RUNNING are ignored by the simulator. We configured TorPS to produce one million circuits using Tor’s standard relay selection algorithm based on the “simple” user model, which emulates a user browsing the web on port 80.

**Distribution of Circuit Risk**

We now examine the riskiness of Tor circuits in our dataset. First, we plot the distribution of circuit risk $\mathcal{U}(g,C)$ for one million circuits generated during the month of December, 2013 in Figure 3. The source and destination of each circuit is chosen uniformly at random from the 249 ISO 3166-1 countries; we do this to present an unbiased assessment of risk regardless of source and destination. Recall that $\mathcal{U}(g,C)$ produces a risk score with respect to a specific adversary country $g$. The “All” line in Figure 3 corresponds to risk of all one million circuits against all 249 potential adversary countries. Overall, 49% of circuits have zero risk, when all possible adversaries $g$ are considered. This result stems from the fact that there are 98 countries which host zero Tor relays and have zero MLATs. In these cases, $M_{g,ci} = 0$ regardless of $c_i$, and thus $\mathcal{U}(g,C) = 0$. This accurately reflects cases where specific countries have no ability to monitor Tor traffic, either directly or indirectly through an ally.

The country-specific lines in Figure 3 correspond to the risk distribution of the one million circuits versus particular adversaries. The US has the highest hostility score ($H_{US} = 0.75$) and the most MLATs (87), so it is not surprising that 71% of circuits in our dataset are at high risk of surveillance $\mathcal{U}(US,C) \geq 0.5$ versus the NSA. A more interesting comparison is between Germany and China: although Germany has lower hostility than China ($H_{DE} = 0.5$ vs. $H_{CN} = 0.66$) and fewer MLATs (37 vs. 56), Figure 3 indicates that
circuits are more at risk to passive attacks from the BND than the MSS. This result is due to the fact that, as of December 2013, Germany hosts the most Tor relays of any country (see Table 1), whereas China hosts zero relays.

![Figure 3: CDF of circuit risk against all adversaries, and against specific adversaries.](image1)

![Figure 4: Change in circuit risk, depending on whether the source/destination are within the adversarial country.](image2)

**Impact of Source and Destination on Circuit Risk**

Figure 3 demonstrates that our circuit risk metric is able to capture the risks posed by adversarial countries with differing Internet surveillance capabilities. Next, we examine the impact of the source and destination countries on circuit risk. Figure 4 plots the CDF of circuit risk when the adversarial country is held constant \(g = \text{US}, \) in this case, while the source and/or destination country is varied. We examine four scenarios: random sources and destinations (the same scenario as is Figure 3), source in the US, destination in the US, and source and destination in the US. The results in Figure 4 are based on the same 1M circuits used in Figure 3.

Figure 4 demonstrates that the circuit risk metric correctly captures the importance of the source and destination in determining overall circuit risk. Random sources and destinations has the lowest risk profile, since many of the chosen countries do not have MLATs with the US (or, are simply not the US). Conversely, when the source and destination are both in the US, risk is uniform and maximized to the US’s country hostility score. The circuit risk metric produces uniform scores in this case because \(H_{CN} = 0.75, \) \(M_{\text{US,US}} = 1, \) and \(\beta \) is weighted towards the first and last connections in each circuit. More importantly, this result is intuitive: deanonymization attacks are most effective against Tor when the adversary can observe the source and destination [13, 15, 41].

The random/US and US/random results fall between the two extremes. Random/US has a lower risk profile than US/random because guard relays tend to have lower risk scores than exits when the US is the adversary. Thus, when destinations are random, the risk of the last connection is dominated by the risk at the exit.
4.3.4 Impact of Circuit Risk

Figure 4 illustrates the complex relationship between sources/guards and exits/destinations. In some cases, the risk of surveillance is dominated by the source and destination, in which case there is nothing Tor can do to reduce risk (e.g. changing to guards and exits in alternate countries, extending the length of the circuit, etc.). However, there are other cases where the risk of surveillance at the source and destination is lower than the risk of surveillance at the guard and exit. In these cases, users’ Tor circuits are more risky because the guard and/or exit are in hostile countries, thus opening up users’ traffic to increased surveillance. In this section, we delve deeper into cases where circuit risk increases due to the selection of guards and exits in hostile countries. These cases are interesting because they represent instances where Tor’s relay selection algorithm could be improved, i.e. by choosing guards and exits in less risky countries. We will examine strategies for reducing risk in § 4.5. As before, our evaluation is based on 1M simulated circuits from TorPS, generated from the December 2013 Tor consensus and directory files.

Distribution of Risk by Participant. We now examine how risk is distributed across the five participants in Tor circuits: the source, guard relay, middle relay, exit relay, and destination. In this experiment, we choose sources and destinations randomly from the set of 249 ISO 3166-1 countries, while the three relays come from our 1M TorPS simulations. Figure 5 presents a boxplot showing Adversary Adjusted Risk ($A_{g,c} = M_{g,c}H_c$) at each position in Tor circuits. The whiskers of each box are the minimum and maximum observed risk, the top and bottom of each box are 5th and 95th percentile risk, and the black bar in each box is median risk. Boxes are shown for four different adversaries. In general, risk tends to be lower for sources and destinations because they are drawn from 249 possible countries. Risk at the guard, middle, and exit relays are higher because our simulations take into account the real-world locations of Tor relays, which are predominantly located in hostile countries (see Figure 2).
The results in Figure 5 confirm our hypothesis that there are many cases when the source and destination countries are less risky than the countries containing the guard and exit relays. For example, in cases where the source is in the adversarial country, risk is maximized (top whisker on the box). Thus, the risk at the guard must be less than or equal to risk at the source. The same observation applies when the destination is in the adversarial country, and you compare risk at the destination and the exit relay. However, 95th percentile risk at the source is only 0.32, while median risk at the guard is between 0.33 and 0.37. This illustrates that for most countries, risk at the source is significantly lower than risk at the guard across four powerful adversaries. Again, the same observation applies when comparing risk at the destination to risk at the exit relay.

**Source Risk versus Guard Risk**

Now that we have confirmed that Tor’s relay selection algorithm sometimes chooses guard and exit relays that increase the risk of surveillance, we now pose three questions. First, how frequently does the Tor client build pathological circuits, e.g. where guard risk is higher than source risk? Second, how many countries are impacted by pathological circuit construction? Third, what fraction of Tor’s users are impacted by pathological circuit construction? To answer these questions, we analyze the difference in ARR between the country containing the guard relay and the source country using our 1M simulated circuits from TorPS. The difference is > 0 if the guard country is more risky, which represents a pathological circuit.

To simplify our analysis, we conservatively assume that the adversary $g = US$, since it is the most powerful adversary in our dataset. Furthermore, we focus on comparing guard and source risk, and fix the destination country as the US (again, as a worst-case scenario). We omit analysis of destinations and exit relays for brevity. Figure 6 plots the difference in ARR between the guard country and source country. The “All” line shows the difference when source countries are selected randomly, and thus presents an overall picture of when guard risk is greater than source risk. We observe that the difference is > 0 (i.e. the guard country is riskier) in 95% of circuits. In 13% of circuits the difference is 0.75, meaning that source risk is zero, versus guard country risk of 0.75 (which occurs when the guard is in the US). The country-specific lines in Figure 6 highlight examples where pathological circuits do and do not exist.

In these experiments, the source country is held constant. When the source is in the US or the UK, source risk is always greater than or equal to guard risk. This is to be expected: the US and the UK are the two most hostile countries in our model (see Table 2). Conversely, when the source is in Japan or Taiwan, 83% and 100% of circuits (respectively) have riskier guards than sources. Overall, our experiments reveal that in 201 countries (out of 249), 50% of circuits have guards in riskier countries than the source. Tor Metrics [2] provides estimates for the number of unique Tor users in each country each
month. According to their user statistics for January 2014, 61% of Tor users reside in countries affected by this pathology. In other words, for the majority of Tor users, the Tor relay selection algorithm is likely to choose a guard node in a country with higher risk of surveillance than the user’s home country (assuming the most powerful adversary, the US).

4.5 Jurisdictional Arbitrage: Novel relay and path selection algorithms and overlays

We have applied the knowledge of MLATs gained to improve the security of networks in our paper on Undersea Cables [144], to reduce the risk posed by attacks from GIAs.

We suggest here other ideas for future work for applying the risk metrics from § 4.3. One simple idea for improving the security of Tor circuits is to exclude relays in risky countries, as defined by § 4.3. Figures 4 and 5 compare the circuit risk C for 1M circuits with all relays allowed, and when relays in particular adversarial countries (US and Germany in these cases) are excluded.

![Figure 17: Circuit risk if relays in the US are disallowed (g = US).](image)

![Figure 18: Circuit risk if relays in Germany and MLAT allied countries are disallowed (g = DE).](image)

Future research will develop new algorithms that build more trustworthy circuits against powerful and complex GIA adversaries, while maintaining performance levels.

Guided Relay Placement- The growth in Tor relays in the past few years has been concentrated in countries which have high hostility scores, as shown in Table 2. It is our hope that our risk metrics will have wide applicability in other security contexts. One interesting direction is to use our metrics to guide the future growth of the anonymous communication networks. Rather than having volunteers place relays in random locations, a concerted effort could be made to install relays in low-risk countries. This could potentially lower the surveillance risk for all users of the system, especially users concerned about observation by GIAs representing major world powers.

Future Work - Incorporating AS Topology. Our current metrics use the geolocation of Tor relays to assess surveillance risk. However, this ignores cases where the flow of
traffic from c1 to c2 flows through other countries that may have greater hostility. In future work, we plan to address this shortcoming by incorporating AS topology into our metrics. This will necessitate measuring the subset of the Internet graph that connects Tor relays, which we anticipate will be challenging.

The results from the first part of the first task give us novel and actionable understanding about the threat of Internet surveillance in different countries, and between allied or at least collaborating countries. However, the gap in these metrics is that they do not account for the underlying AS topology of the Internet. Specifically, a GIA could facilitate passive surveillance by monitoring traffic routing through ASes under the GIA's control. In this way, a GIA in country g could monitor traffic flowing through ASes that are not within g or countries allied with g. See Figure 2: Hostility of countries versus the number of Tor relays in that country.

**Leveraging Hostility Scores for Tor Circuit Construction**

If users of anonymous networks face threats from lawful interception by countries through which their network traffic passes, it follows that those threats may be mitigated by rerouting through countries with greater trust, defined as less legal and government hostility to that traffic. The hostility index may be incorporated into Tor's path selection algorithm to select relays and bridges in countries with low hostility indices in relation to the user's geographic location. More research needs to be done testing the use of the hostility index in actual path selection.

**Challenges**

A challenge for this research task is constructing a sufficiently accurate map of the Internet. Many studies have attempted to reconstruct the Internet’s AS topology using data from iPlane [48], CAIDA traceroutes [23], Looking Glass servers, RouteViews [4], PlanetLab [58, 59] and others [3]. We have begun to leverage these data sources, and plan to continue this work. We have comprehensively analyzed the Tor network at an AS-level, and created an AS-level map of the Tor relay network, both of which are crucial to the success of the research. We leave for future work to leverage the measurement infrastructure developed earlier, enabling reconstruction of the AS topology around existing Tor relays.

**5 Related Research Projects**

**5.1 MLATs and the Undersea Cable Network**

This work builds on the MLAT and Jurisdictional arbitrage work, applying it to undersea cables. [144, 152].
5.2 MLATs and Legalizing Domestic Surveillance: The Role of Mutual Legal Assistance Treaties in Deanonymizing TorBrowser Technology

This paper presents the first analysis of the impact of MLATs on network security in actual case law, and how CSPs like Google and Microsoft have become deeply involved behind the scenes in MLAT implementation in communications infrastructure for dynamic triggering of surveillance. Key findings include that MLATs can represent legal “devil’s bargains” with cartel partner countries. For example, the US has MLATs with several countries who regularly prosecute apostasy with death sentences. Adding constitutional protections would strengthen MLATs. And that recent “reforms” are funded by major corporations like Google and Facebook to streamline their operations in foreign countries. This work is presented in the paper Legalizing Domestic Surveillance: The Role of Mutual Legal Assistance Treaties in Deanonymizing TorBrowser Technology [145].

5.3 MLAT.is: Web portal with network database query and visualization capabilities

MLAT.is is a scalable service providing accurate predictions of treaty partner cooperating jurisdictions for internet routing and path selection algorithms. MLAT.is builds a structural model of all world jurisdictions. We construct an annotated map of world jurisdictions by treaty partners. Users may query the MLAT database and return clusters of treaties, as well as their relationship visualizations. We relate MLATs to the Internet and review their impact. We predict risk impact by composing measured performance of clusters of known MLA treaty partners. This method allows us to accurately and efficiently predict the effects of MLAs on path selection risk between arbitrary Internet hosts. The MLAT.is service has been access over 17,000 times as of August, 2017 by over 40 countries.

5.4 MLAT Graph: Measuring the worldwide MLATs country cartel graph

MLATs reveal otherwise invisible cartels of cooperating countries connected by international law enforcement treaties. This involves applying network theory measurements such as centrality to investigate whether they can reveal traffic concentrations in specific legal jurisdictions. For example, centrality may reveal Tor traffic concentration in certain ASes, which may help guide path selection algorithms to improve network security, anonymity and privacy. This work is published in MLAT Overlay [146]. This work is covered under IRB #17-04-16-MLATs and the Darknet, which was approved April 22, 2017.

5.5 Dynamic Triggering: MLATs in Lawful Intercept for communications networks

This is another project which builds on the dissertation work. This project includes two subtasks. The first involves researching the application of MLATs in case law to determine their actual impact. The second involves analyzing the involvement of major
CSPs like Google and Microsoft in MLAT implementing *dynamic triggering* of lawful intercept in CSPs.

**MLAT Case Law**

We have performed an extensive MLAT case law search and analysis in the *Richmond JOLT* paper [145]. We continue our review of recent decisions, including the September 30, 2014 and related Second Circuit rulings in *US v. Getto*, and the September 2013 First Circuit rulings in *Boston College Trustees*, among other case law.

**Dynamic Triggering, MLATs, Google and Microsoft**

In addition to being a possible trust factor for lawful intercept, MLATs provide their own attack vector for anonymous communications networks. MLATs enable legal pressure on CSPs worldwide, and so offer a significant vector to attempt to “break” Tor or its underlying encryption. Legal pressure on CSPs like Facebook and Google by foreign governments to locate surveillance-enabled servers in their countries, has led them to seek to incorporate into MLATs the automation of remote global surveillance. Their industry groups, like the International Chamber of Commerce (ICC), seek to incorporate technical standards for automated surveillance LI into MLATs, through remote “dynamic triggering.” [128, 129] A multi-stakeholder group, the Global Network Initiative, funded principally by Microsoft and Google [130, 131], announced its agenda on January 28, 2015 in “Data Beyond Borders: Mutual Legal Assistance in the Internet Era.” In it, GNI sets forth a strategy a to further shape MLAT, and thus, surveillance policy [127]. GNI states:

[We] recommen[d] that LEAs and governments...improve existing MLATs so that...they cover evolving IP-based communications services...includ[ing]...interception of electronic communications,...both communications content and communications data...” [127- 131]

This separate research makes the connection between major US-based CSPs, MLAT policy, and surveillance. It investigates and analyzes privacy implications resulting from delineating and comparing Fourth Amendment search and seizure endpoint options in communications technology like cellphones, skype, VoIP, chat, and streaming chat. This work is in draft.

**5.6 Jurisdiction of the Darknet: Successful LEA De-Anonymization**

This project researches eight major darknet investigations and the methods by which users were successfully deanonymized by LEAs. Observations and findings include that key darknet investigations reveal advances in LEA deanonymization capabilities. Also, that NITs have held up under legal scrutiny but create issues of congruence with the CFAA. And, that expert witness testimony has yet to establish whether NITs meet the Daubert legal standard. This work is in draft.
5.7 CircuitBlasTor: Measuring The Tor Graph for Randomness

In a separate project, we have built on the Jurisdictional Arbitrage [147] work and MLAT.is database model the elements of the Tor graph, to better understand circuits and traffic concentration. We briefly describe this here, however it an extension of the original project. This work is covered under IRB #17-04-17-CircuitBlasTor, Tor circuits, which was approved April 22, 2017.

The “Tor Internet Graph” consists of nodes and undirected (bidirectional) edges. Nodes include not only the Tor relays themselves, but also the routers in Autonomous Systems (ASes), and Internet exchange points (IXPs or IXs), all of which carry Tor traffic. Each node is an IP, so it is also a graph of IPs, and each node also lies in an AS, so it is also a graph of ASes. The edges are links connecting IPs in different ASes.

Tor is an overlay network of the Internet. In understanding its vulnerabilities to threats and attacks, it is useful to apply tools of network science to better understand it. Today’s Internet consists of $2^{32}$ or about 4.2 billion theoretically possible IPv4 addresses. Total paths between all IP would include, then, approximately $4.2^2 / 2$ billion possible paths, without even considering IP or AS growth and change rates. Assuming an average of 32 hops between hosts, I might end up considering a dataset of 564.48 sextillion ($5.6448 \times 10^{20}$) or $282,240,000,000,000,000,000$ hops, give or take a few quintillion. No source provides this data, to our knowledge. Further, it seems this level of data might provide needlessly precise path information. And, IPv6 addresses occupy an even greater space.

In this research task, I inquire as to where a global passive adversary who focuses surveillance on the Tor anonymous network might focus and maximize their resources and results. And, whether I can make recommendations regarding future relay placement or routing to diversify traffic concentration and decrease this vulnerability? Specifically, our areas of inquiry include centrality of nodes by AS. What, if anything, do centrality results and analyses tell us about the vulnerability of Tor to adversaries seeking concentrations of Tor traffic? How do these results differ from frequency analyses? How do they differ from traffic between a control group of semi-random non-Tor nodes? And, with respect to node recommendations, can I model where future relays are best placed to diversify routes and decrease traffic concentration globally across the Internet?

Past research has investigated the ASes in which Tor relays themselves lie, or in which Tor client nodes or destinations lie. Our research focuses not on ASes of Tor nodes per se, but on ASes through which Tor traffic passes. Thus, I characterize the former set of inquiries as concerning node location concentration analysis, and distinguish our inquiry as concerning Tor traffic location concentration analysis.

Network Science provides the tools for performing this analysis, and, indeed, graph analysis has been applied to the Internet [93, 94, 95, 96, 97, 126]. R and the Python libraries provided by NetworkX are two tools which can enable our analysis [133, 140].

To determine whether and where Tor traffic may concentrate, I design a series of experiments. The first two experiments would be designed to detect and identify concentrations of Tor ASes and therefore Tor traffic.

The first experiment measures AS frequency. This experiment, using tracerouted data, examines individual hops. We determine the AS of each hop using Maxmind
geoiplookup database [114], based on the specific IP (full IP). We then sort by frequency of AS, to determine what ASes might display significant concentrations of Tor traffic. We perform this experiment on both traceroutes collected from 63 PlanetLab servers, and tracerouted data from 176 iplane servers. We hypothesize that ASes that occur most frequently could indicate ASes where maximum Tor traffic is concentrated, and thus where an adversary’s resources and focus could profitably be allocated. We experiment with a control group of all traffic, including Tor & non-Tor, in the overall iplane Internet network. We hypothesize that Tor network results would differ from those of the semi-random iplane network, to test the validity of our methods of identifying concentrated Tor traffic.

The second experiment measures Tor graph centrality. For this experiment, I convert all individual traceroute records into [ip, ip] pairs. Taking all hop adjacencies and creating one [ip, ip] pair for every ip adjacency set would be one method to achieve this. We could then easily convert [ip, ip] pairs into [as,as] pairs to create the Tor Internet graph, as a graph of ASes. We hypothesize that a high centrality could also indicate that Tor traffic was flowing through certain ASes at higher rates than others. This would support the central hypothesis, that “high-value” locations exist where an adversary could best concentrate surveillance resources.

Third, I experiment with a series of graph metrics, examining the following:

1) Nodes- can I confirm our AS dataset sample size?
2) Links- how many connections exist between AS nodes?
3) Symmetric links - how many AS pairs have matching pairs in the other direction?
4) Average outdegree - to how many other ASes is each AS connected, classified by outbound? Do in- and out-degree match, as would be expected in a complete graph?
5) Average indegree- to how many other ASes is each AS connected, classified by inbound?
6) Clustering coefficient - the degree to which ASes tend to cluster together?
7) Assortativity (in/in): the tendency of ASes to connect to other ASes of similar degree, (in/in)?
8) Assortativity (out/out): the tendency of ASes to connect to other ASes of similar degree, (out/out)
9) Radius - average shortest path between ASes?
10) Diameter - maximum eccentricity- greatest distance between any pair of ASes?
11) Average path length - average number of ASes per path overall?
12) Connected components - are AS subgraphs present?

We hypothesize the implications of the metrics or frame the third experiment’s queries as follows:

1) Nodes- that our AS dataset sample size is valid
2) Links- does a higher proportion of links to nodes reveal anything about AS concentration?
3) Symmetric links - whether fewer symmetric links inform about the unidirectional flow of traffic, and thus inform about recommendations for future node placement to diversify ASes and therefore traffic flow?
4) Average outdegree - does a higher outdegree indicate opportunities to diversify routing and thus traffic concentration?
5) Average indegree- does a higher indegree indicate opportunities to diversify routing and thus traffic concentration?
6) Clustering coefficient - does a high clustering coefficient indicate an opportunity to diversify routing and thus traffic concentration?
7) Assortativity (in/in): do ASes with high indegrees tend to connect to other ASes with high indegrees? If so, is that related to traffic concentration?
8) Assortativity (out/out): do ASes with high outdegrees tend to connect to other ASes with high outdegrees? If so, is that related to traffic concentration?
9) Radius - how many AS hops does traffic have along shortest paths, in general? Does that represent opportunities to diversify routing paths and thus, traffic concentration?
10) Diameter - maximum eccentricity- does eccentricity reveal paths with less traffic concentration?
11) Average path length - how does this compare to radius? How many AS hops does traffic have, in general? Does that represent opportunities to diversify routing paths and thus, traffic concentration?
12) Connected components - Would the presence of one or more subgraphs lessen the likelihood of traffic concentration? Or, present greater challenges when allocating surveillance resources?

The results are presented in another paper, CircuitBlasTor, which has been submitted for publication.

5.8 Diameter of the Darknet-hyperlink analysis

This is yet another separate project, which builds upon the work in the dissertation. Based upon the largest darknet dataset so far, we map and measure the darknet, visualize it and make observations and measurements, comparing it to the surface web. We are collaborating with University of Portsmouth, UK researchers on this work, which is in draft. This work is covered under IRB #17-04-18-Diameter of the Darknet, which was approved April 22, 2017.

References


[19] Richard Chirgwin, The Register, Australia the idiot in the global village, says Geoff Huston, 2 Sep 2015, http://www.theregister.co.uk/2015/09/02/australia_the_idiot_in_the_global_village_says_geoff_huston


[30] e. g. CanaryWatch, https://canarywatch.org/about.html
http://www.apple.com/pr/pdf/131105reportongovinforequests3.pdf, 

https://www.google.com/transparencyreport.


[37] Manuel Egele, Peter Wurzinger, Christopher Kruegel, and Engin Kirda, Defending Browsers against Drive-by Downloads: Mitigating Heap-spraying Code Injection Attacks, Sixth Conference on Detection of Intrusions and Malware and Vulnerability Assessment (DIMVA), Milan, Italy, June 2009


[88] Anomaly-based censorship-detection system for Tor,
https://research.torproject.org/techreports/detector-2011-09-09.pdf and
https://metrics.torproject.org/userstats-censorship-events.html

[89] Open Rights Group,


[91] The Arbeitskreis Vorratsdatenspeicherung (German Working Group on Data Retention), http://wiki.vorratsdatenspeicherung.de


[93] AL Barabási, R Albert, Emergence of scaling in random networks, Science 286 (5439), 509-512, 1999

[94] Scale-free characteristics of random networks: the topology of the world-wide web

[95] Modeling the Internet's large-scale topology, SH Yook, H Jeong, AL Barabási - Proceedings of the National Academy of Sciences, 2002 99 (21), 13382-13386


[98] ETSI TS 102 677 V0.4.1 (2010-03), Lawful Interception (LI); Dynamic Triggering:
Dynamic Triggering of Content of Communication Interception, Work Item:
DTS/LI-00058, 3rd[sic] GENERATION PARTNERSHIP PROJECT (March, 2010)
[hereinafter ETSI TS 102 677],


[111] The Tor path simulator (TorPS). https://torps.github.io


[113] Greg’s Cable Map, cablemap.info.

[114] Maxmind geoip database


[121] EFF, Mandatory data retention, Argentina https://www.eff.org/issues/mandatory-data-retention/argentina


[125] OECD, Mutual Legal Assistance
Arrangements within Asia-Pacific,
https://www.oecd.org/site/adboecdanti-corruptioninitiative/40046104.pdf


[130] Guidestar.org, Global Network Initiative (GNI)
//www.guidestar.org/organizations/27-2322782/global-network-initiative.aspx,

[131] IRS, Global Network Initiative (GNI), Form 990, GNI, 2013

http://www.inf.uni-konstanz.de/algo/publications/b-fabc-01.pdf

[133] NetworkX: High-productivity software for complex networks,
https://networkx.github.io

[134] CAIDA.org, IPV4 and IPV6 AS core, AS-Level Internet Graph
http://www.caida.org/research/topology/as_core_network/pics/ascore-2011-apr-ipv4v6-standalone-1600x876.png


[140] R, statistical computing and graphics, R version 3.2.2 (Fire Safety) https://www.r-project.org


61


[150] Sarah Cortes, CircuitBlasTor, submitted.

[151] Sarah Cortes, Andrew Lewman, Gareth Owen, Diameter of the Darknet, graph analysis of Tor Hidden Services, in draft.

[152] Sarah Cortes, MLAT Database, available through MLAT.is web portal.

[153] Sarah Cortes, MLAT.is web portal.

Appendix A- Contributions

By Contribution to Knowledge
(Related research artifacts)

● The first real-world schema for measuring node LI trust worldwide on a country level, Jurisdictional Arbitrage: Quantifying and Counteracting the Threat of Government Intelligence Agencies Against Tor [147]:
   a. Identification of five lawful interception (LI) risk factors
   b. Data collection on those factors
   c. Definition of Risk Metrics utilizing those factors and data

● The first publicly available comprehensive global MLAT database [152]
● The first application of aspects of that schema, to the undersea cable network [144]
● The first analysis of legal impact of MLATs on LI and public policy [145]
• Implementation into actual use via MLAT.is tool of the MLAT database web portal with database query and network visualization features [153]

• Provides a foundation for several subsequent research projects in various stages of grant proposals, publication, submission or draft:
  
  a. **MLAT World Treaty Cartel Internet Overlay**: Analysis of the MLAT country cartel network itself: the first application of graph theory to MLATs, and how results may impact network communications traffic location [146].

  b. **Dynamic Triggering**: ETSI MLAT Lawful Interception of communications networks, The first analysis of the impact of MLATs on network security in actual case law, and how CSPs like Google and Microsoft have become deeply involved behind the scenes in MLAT implementation in communications infrastructure for dynamic triggering of surveillance [148].

  c. **Jurisdiction of the Darknet: Successful LEA De-Anonymization**: The first comprehensive legal and technical analysis of successful deanonymization by LEAs. Balancing approaches to investigations [149].

  d. **CircuitBlasTor**: An analysis of Tor vulnerabilities revealed by the first application of graph theory to the public Tor network [150].

  e. **Diameter of the Darknet** - the first hyperlink analysis of the Darknet based on a unique and comprehensive dataset gleaned from Hidden Services [151].

**By Research Artifacts**

Papers (published):

1. **Jurisdictional Arbitrage-USENIX** and **NSF submitted- in revision** [147]

2. **Undersea Cables-PoPETS 2015** [144]

3. **MLATs and LI-Richmond JOLT 2015** [145]

4. **MLAT Internet Overlay-Digital Traffic Analytics-IEEE-HST** [146]

5. **Dynamic Triggering: ETSI MLAT LI in comm networks- draft** [148]

6. **Jurisdiction of the Darknet-Success LEA De-Anonymization-draft** [149]

7. **CircuitBlasTor-Computers & Society**, submitted, in revision [150]

8. **Diameter of the Darknet- in draft** [151]

Tools/Applications (implemented):

9. **MLAT database** [152]

10. **MLAT.is** web portal, database query and visualizations [153]

**By Scientific or Academic Field**

**Computer Science- Information Assurance- Privacy, Network Security**

1. **Jurisdictional Arbitrage-USENIX** and **NSF submitted- in revision**

2. **Undersea Cables-PoPETS 2015**

3. **Dynamic Triggering: ETSI MLAT LI in comm networks- draft**

4. **MLAT database**

5. **MLAT.is** web portal and visualizations

**Network Science**

6. **MLAT Internet Overlay-Digital Traffic Analytics-IEEE-HST**
By Key Findings and Observations

Jurisdictional Arbitrage [147]

- 71% of circuits in our dataset are at high risk of surveillance by the US.
- Although Germany has lower hostility than China (HDE = 0.5 vs. HCN = 0.66) and fewer MLATs (37 vs. 56), circuits are more at risk to passive attacks from the BND than the MSS. This result is due to the fact that, as of the experiment date, Germany hosts the most Tor relays of any country, whereas China hosts zero relays.
- When destinations are random, the risk of the last connection is dominated by the risk at the exit.

![Figure 4: Change in circuit risk, depending on whether the source/destination are within the adversarial country](image)

- There are many cases when the source and destination countries are less risky than the countries containing the guard and exit relays
- Thus, the risk at the guard must be less than or equal to risk at the source.
- 95th percentile risk at the source is only 0.32, while median risk at the guard is between 0.33 and 0.37. This illustrates that for most countries, risk at the source is significantly lower than risk at the guard across four powerful adversaries
- We have confirmed that Tor’s relay selection algorithm sometimes chooses guard and exit relays that increase the risk of surveillance

Considering these results, the following key questions arise:

- How frequently does the Tor client build pathological circuits, e.g. where guard risk is higher than source risk?
- How many countries are impacted by pathological circuit construction?
- What fraction of Tor’s users are impacted by pathological circuit construction?

Observations and conclusions:
• We observe that the difference is > 0 (i.e. the guard country is riskier) in 95% of circuits. In 13% of circuits the difference is 0.75, meaning that source risk is zero, versus guard country risk of 0.75 (which occurs when the guard is in the US).
• When the source is in the US or the UK, source risk is always greater than or equal to guard risk. This is to be expected: the US and the UK are the two most hostile countries in our model. Conversely, when the source is in Japan or Taiwan, 83% and 100% of circuits (respectively) have riskier guards than sources.
• Our experiments reveal that in 201 countries (out of 249) 50% of circuits have guards in riskier countries than the source. Tor Metrics [2] provides estimates for the number of unique Tor users in each country each month. According to their user statistics for January 2014, 61% of Tor users reside in countries affected by this pathology. In other words, for the majority of Tor users, the Tor relay selection algorithm is likely to choose a guard node in a country with higher risk of surveillance than the user’s home country (assuming the most powerful adversary, the US).

Undersea Cables [144]
• In all cases, MLATs can allow for much greater reach by a hypothetical adversary country seeking to compromise a path and deanonymize a user.

MLATs and LI [145]
• MLATs can represent legal “devil’s bargains” with cartel partner countries
• The US has MLATs with several countries who regularly prosecute apostasy with death sentences
• Adding constitutional protections would strengthen MLATs
• Recent “Reforms” are funded by major corporations like Google and Facebook to streamline their operations in foreign countries

MLAT Internet Overlay-Digital Traffic Analytics [146]
• Countries that enter into MLATs cluster themselves into cooperating cartels
• Network science analysis of MLAT relationships can help reveal these clusters, perhaps helping predict the next country to join each cluster

Dynamic Triggering: ETSI MLAT Lawful Intercept comm networks [148]
• Privacy implications resulting from delineating and comparing Fourth Amendment search and seizure endpoint options in communications technology like cellphones, skype, VoIP, chat, and streaming chat.

Jurisdiction of the Darknet: Successful LEA De-Anonymization [149]
• Legal and technical analysis of successful deanonymization by LEAs
• Key darknet investigations reveal advances in LEA deanonymization capabilities
• NITs have held up under legal scrutiny but create issues of congruence with the CFAA
• Expert witness testimony has yet to establish whether NITs meet the Daubert legal standard
CircuitBlasTor [150]
- Tor paths form a complex temporal network
- Discrepancies in randomness of Tor relays may cause vulnerabilities
- Tool that creates real Tor circuits rather than simulated circuits

Diameter of the Darknet [151]
- The giant component in the darknet is small but growing
- Largest known database of Tor Hidden Services data

MLAT database [152]
- Number of worldwide bilateral MLATs exceeds 1,000
- Even nation-states like Iran have countries willing to enter into MLATs

MLAT.is web portal, database query and visualizations tool [153]
- Over 17,000 accesses since publication by users in over 40 countries

Appendix B- Research Deployment and Administration

Broader Impact Through Deployment
The research outlined in this proposal has the potential for large-scale impact. If the efforts outlined here are successful, the results immediately improve safety and security for all current and future Tor users. In addition, the research provides public policy recommendations to strengthen and balance MLATs. This has the potential to impact thousands of investigations and prosecutions in countries worldwide, by providing a more balanced view of the impact, better LEA investigations to catch criminals on tor, as well as helping tor users better assess the risks of their circuits.

The metrics and algorithms developed as part of the research can have immediate impact through deployment. All of the datasets, algorithms, and code are open-sourced (see Data Management plan) in the hope that they will be useful to a wide-range of Internet security systems/researchers. By collaborating closely with Tor, I have also laid the groundwork to facilitate deployment of the technologies as part of Tor’s suite of tools, with a positive impact on the millions of regular Tor users.

Collaboration
In addition to my NEU colleagues, I have collaborated directly with current and former Tor staff and affiliates in all stages of the proposals and research. I also collaborated with US Naval Research Laboratory (NRL) staff. I collaborate with ETSI staff in researching relevant technology standards related to the impact of MLATs on dynamic triggering of Lawful Intercept (LI). And, I collaborate with researchers at the University of Portsmouth, UK in the Hidden Services research.
Data Management Plan
We are committed to public access for the products of this research. We have already made known in the technology, legal and public policy communities the availability of this data. We aim to continue on this path for the research output of this proposal. Output of the research includes (a) data sets from traces and measurement studies, (b) metrics and software that implements these metrics for quantifying risk to anonymous systems, and (c) tools like MLAT.is and websites that implement visualizations and provide protection mechanisms. In this section, I describe plans for collection, organizing, managing, and releasing these products.

Data retention
We will retain the raw data, summarized and reported data, as well as all publications and software for at least one decade. All software, code, scripts and raw data are be stored on research machines. Remote access to server resources are provided through SSH and other industry standard remote-access tools.

Data format and analysis tools
All released data includes documentation that describes how the data was collected and the format in which the data is represented. Typically open-source-compatible formats are used (ASCII, CSV, etc.), used for research data for this project as well. All code, scripts, and metadata that I use as a part of conducting the research have been released alongside the data itself to authorized parties. This allows other authorized researchers to replicate our results (using either our released data, or data set that they collect on their own), and also easily enables further research on our released data by other groups.

Data dissemination
The software and publicly released data sets are served from my Web server(s), alongside my research servers.

Appendix C- Committee Background and Justification
Prof. Albert-László Barabási is the Robert Gray Dodge Professor of Network Science and a Distinguished University Professor at Northeastern University, where he directs the Center for Complex Network Research, and holds appointments in the Departments of Physics and College of Computer and Information Science (CCIS), as well as in the Department of Medicine at Harvard Medical School and Brigham and Women’s Hospital in the Channing Division of Network Science, and is a member of the Center for Cancer Systems Biology at Dana Farber Cancer Institute. He received his Masters in Theoretical Physics at the Eötvös University in Budapest, Hungary, and his Ph.D. three years later at Boston University. Barabási has authored several books and hundreds of articles. His work lead to the discovery of scale-free networks in 1999, and proposed the
Barabasi-Albert model to explain their widespread emergence in natural, technological and social systems, from the cellular telephone to the WWW or online communities.

Barabási is a Fellow of the American Physical Society. In 2005 he was awarded the FEBS Anniversary Prize for Systems Biology and in 2006 the John von Neumann Medal by the John von Neumann Computer Society from Hungary, for outstanding achievements in computer-related science and technology. In 2004 he was elected into the Hungarian Academy of Sciences and in 2007 into the Academia Europaea. He received the C&C Prize from the NEC C&C Foundation in 2008. In 2009 APS chose him Outstanding Referee and the US National Academies of Sciences awarded him the 2009 Cozzarelli Prize. In 2011 Barabási was awarded the Lagrange Prize-CRT Foundation for his contributions to complex systems, awarded Doctor Honoris Causa from Universidad Politécnica de Madrid, became an elected fellow in AAAS (Physics), then in 2013 Fellow of the Massachusetts Academy of Sciences and, just recently, the 2014 Prima Primissima Award for his contributions to network science by the Hungarian Association of Entrepreneurs and Employers.

Prof. Thomas H. Koenig’s primary fields of research specialization include International Law; Internet Policy; Law and Society; Informatics; and Criminology. He has been a faculty member at Northeastern University since 1977, serving 2002-8 as Chair of the Department of Sociology and Anthropology. Formerly a Liberal Arts Fellow in Law and Sociology at Harvard University Law School, he has served as a visiting professor at Tufts University, Brown University, SUNY at Buffalo and in Hungary’s program on Post-Soviet Change Management. He has testified before both Houses of Congress and his research has been discussed in the Economist, Wall Street Journal, New York Times, Time, U.S. News and World Report, as well as being cited thousands of times in leading academic publications and in hundreds of court decisions. Professor Thomas Koenig specializes in Internet law and governance. He has been involved with the Internet since it was the ARPANET and wrote his Ph.D. dissertation using a social network algorithm to map interconnections among America’s top business leaders. Professor Koenig’s research focuses on both empirical and theoretical examination of, among other areas, online privacy, the degree of allowable internet anonymity, and cross-national conflicts over Internet laws.

Guevara Noubir holds a PhD in Computer Science from the Swiss Federal Institute of Technology in Lausanne (EPFL) (1996). His research covers both theoretical and practical aspects of privacy, security, and robustness in networked systems. Prior to joining Northeastern University, he was a senior researcher at CSEM SA (1997-2000) where he led the design and development of the data protocol-stack of the third generation Universal Mobile Telecommunication System (UMTS) and its world first 3G prototype. His research led to a wide range of mechanisms and algorithms for scalable, secure, private, and robust wireless and mobile communications. He led the winning team of the 2013 DARPA Spectrum Cooperative Challenge against 90 academic and industry teams. He is a recipient of the National Science Foundation CAREER Award (2005), the ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec) best paper
award in 2011 and runner-up best paper in 2013. His research was featured in the NSF CISE/CNS Highlights in 2009 and 2012. Professor Noubir has held visiting research positions at Eurecom, MIT, and UNL. Professor Noubir has served as program co-chair of many conferences in his areas of expertise, including the ACM Conference on Security and Privacy in Wireless and Mobile Networks, IEEE Conference on Communications and Network Security, and IEEE WoWMoM. He also co-chaired two NSF Workshop on bio-computation and communications. He serves on the editorial board of the IEEE Transaction on Mobile Computing, the Elsevier Journal on Computer Networks and the ACM Transactions on Information and System Security.

Paul Syverson, PhD, an Association for Computing Machinery (ACM) Fellow, is inventor of Onion Routing, for which he received the Edison Invention Award, and designer of all three generations of Onion Routing systems, including the latest system, Tor. Tor is the largest deployed and used network of its kind in existence (2.5m users daily and over 6000 network servers worldwide). Dr. Syverson has been designing and analyzing security and privacy systems as a mathematician at the Naval Research Laboratory Center for High Assurance Computer Systems since 1989. He has been chair of eight conferences and workshops ranging from the European Symposium on Research in Computer Security to the Privacy Enhancing Technologies Workshop and the Financial Cryptography Conference. He is the editor of several books on these topics, as well as author of many dozens of papers published in refereed conferences and journals. He is also the author of Logic, Convention, and Common Knowledge, a book that discusses philosophical foundations of logic, and employs game theory and distributed computing.

He is former editor of IEEE Cipher. He has been an invited visitor at the Newton Institute for Mathematical Sciences in Cambridge England and was on the faculty of the first International School on Foundations of Security Analysis and Design in Bertinoro Italy. Degrees: PhD and MA in philosophy (logic), MA in mathematics (all three from Indiana), AB in philosophy from Cornell. Syverson has authored more than 100 publications. He holds three patents, including the patent for "Onion routing network for securely moving data through communication networks," issued in 2001. His awards include the Edison Invention Award, for invention of Onion Routing, NRL, 2001; Service Award, Association for Computing Machinery, 2008; Award for Projects of Social Benefit, Free Software Foundation, 2010; Pioneer Award, Electronic Frontier Foundation, 2012; Top 100 Global Thinkers, Foreign Policy, 2012; and the Test of Time Award, USENIX Security Symposium, 2014.

Previous Committee Member
Prof. Engin Kirda- In addition to being co-founder and chief architect at Lastline, a network-based security breach detection firm, Engin Kirda is a Professor at Northeastern University, and the director of the Northeastern Information Assurance Institute. Before that, he has held faculty positions at Institut Eurecom in the French Riviera and the Technical University of Vienna where he co-founded the Secure Systems Lab that is now distributed over five institutions in Europe and US. Engin's recent research has focused on
malware analysis (e.g., Anubis, Exposure, Fire) and detection, web application security, and practical aspects of social networking security. His recent work on the deanonymization of social network users received wide media coverage. He has co-authored more than 90 peer-reviewed scholarly publications and served on program committees of numerous well-known international conferences and workshops. In 2009, Engin was the Program Chair of the International Symposium on Recent Advances in Intrusion Detection (RAID), in 2010/11, Program Chair of the European Workshop on Systems Security (Eurosec), and the Program Chair of the well-known USENIX Workshop on Large Scale Exploits and Emergent Threats in 2012. In the past, Engin has consulted with the European Commission on emerging threats, and recently gave a Congressional Briefing in Washington D.C. on advanced malware attacks and cyber-security.
THESIS TITLE: Jurisdictional Arbitrage: Quantifying and Counteracting the Threat of Government Intelligence Agencies Against Tor

AUTHOR: Sarah L. Cortes

Ph.D. Thesis Approved to complete all degree requirements for the Ph.D. Degree in Computer Science- Information Assurance.

Thesis Advisor-Prof. Thomas H. Koenig

Thesis Reader-Prof. László Barabási

Thesis Reader-Prof. Guevara Noubir

Thesis Reader-Paul Syverson, Ph.D

GRADUATE SCHOOL APPROVAL:

Director, Graduate School

COPY RECEIVED IN GRADUATE SCHOOL OFFICE:

Recipient's Signature

Distribution: Once completed, this form should be scanned and attached to the front of the electronic dissertation document (page 1). An electronic version of the document can then be uploaded to the Northeastern University-UMI website:

http://library.northeastern.edu/get-help/theses-dissertations/submit-your-thesis-or-dissertation

https://secure.etdadmin.com/cgi-bin/activateacct?accountId=2126733;siteId=282;tool=student

https://search-proquest-com.ezproxy.neu.edu/browseresults/734EFFC6AE984EB3PQ/1?accountid=12826
Graphics

Figure 19. Diameter of the Darknet: Minimum Spanning Tree Visualization

Figure 22. Global MLAT cartel network visualization

IRBs

This work includes research covered by the following approved Northeastern University IRBs:

1) IRB #17-04-16-MLATs and the Darknet-Approved 4/22/17.
2) IRB #17-04-17-CircuitBlasTor-Tor circuits-Approved 4/22/17.

Figure 20. Diameter of the Darknet: Clustering visualization

Figure 21: Diameter of the Darknet: thehub7gqe43miyc.onion visualization
Figure 23. Greg’s Cable Map [113] of Undersea Communications Cables

Figure 24: Dynamic Triggering of Lawful Intercept, Single Operator Common MF/CCTF Model, ETSI 102 677