Charting Censorship Resilience and Global Internet Reachability: A Quantitative Approach

Marina Ivanović
ETH Zurich

François Wirz
ETH Zurich

Jordi Subirà Nieto
ETH Zurich

Adrian Perrig
ETH Zurich

***Abstract*—Internet censorship and global Internet reachability are prevalent topics of today's Internet. Nonetheless, the impact of network topology and Internet architecture to these aspects of the Internet is under-explored. With the goal of informing policy discussions with an objective basis, we present an approach for evaluating both censorship resilience and global Internet reachability using quantitative network metrics, which are applicable to current BGP/IP networks and also to alternative Internet network architectures. We devise and instantiate the metric on the network topology of multiple countries, comparing the BGP/IP network, an overlay network using a waypoint mechanism for circumventing undesired nodes, and the path-aware Internet architecture SCION. The novelty of the approach resides in providing a metric enabling the analysis of these aspects of the Internet at the routing level, taking into account the innate properties of the routing protocol and architecture. We demonstrate that the Internet topology matters, and strongly influences both censorship resilience and reachability to the global Internet. Finally, we argue that access to multiple paths accompanied with path-awareness could enable a higher level of censorship resilience compared to the current Internet, and reduce the centralization of Internet routing.**

***Index Terms*—quantitative metrics, censorship, networking, routing, reachability, next-generation Internet architectures**

I. Introduction

The issue of Internet censorship—the deliberate restriction or suppression of information [1], [2]—has emerged as a pervasive concern in the digital era. There have been long standing records of censored network communication practices employed by various entities, and most prominently governments [3]–[7]. Furthermore, the issue of dependency on certain countries has also grown in the context of the global Internet [8], [9], with various analyses that western countries have gained significant influence on the global Internet routing, as routing paths predominantly traverse them [10].

The innate properties of Internet topologies and architectures play a crucial role in determining how traffic flows through the network, and therefore, they are likely to influence the effectiveness of censorship efforts, and in general Internet reachability. In the context of the global Internet, we refer to the network topology as the interconnectedness of Autonomous Systems (ASes), while the Internet architecture encompasses the underlying structure and protocols that facilitate operation of the Internet. For instance, the core Internet routing protocol is the Border Gateway Protocol (BGP), which provides Internet inter-domain routing [11]. In traditional networks using BGP, ASes only consider the next hop when making routing decisions. Unlike traditional routing, SCION—a next-generation Internet architecture designed to provide secure inter-domain routing [12]—ensures that packets traverse predetermined paths and making end-nodes in the network path-aware [12]. Finally, the usage of Virtual Private Network (VPNs) has been a popular technique for Internet censorship evasion, given that it could not only provide an additional layer of secrecy using encryption, but also circumvent censoring devices altogether [13], [14].

Previous research underlines the evidence that the topology of the network could be an indicator of deployed censorship capabilities [15]–[17], and reachability to the global Internet [9], [10], [18]. Nonetheless, to the best of our knowledge, it is an open research challenge how traditional BGP routing, the use of waypoint network with VPN nodes, and in general fundamentally different approaches such as BGP and SCION could be quantitatively compared in this context.

Research Question. In this context, the following research questions arise: *Do the topology and the architecture of the Internet have an influence on Internet censorship, and in general global Internet reachability?* And if so, *can we quantify this influence?* Answering these questions does not only provide insights into the interplay between Internet topology and architecture, and censorship and reachability, but can also quantitatively inform policy-makers. To achieve this, we propose a concrete approach for evaluating censorship and reachability aspects using a quantitative network metric.

Key Contributions.

1) We design a quantitative metric instantiable to *censorship resilience* and *global Internet reachability*. The metric is agnostic to network topology, and applicable to the current Internet and captures path-awareness. (Section III).
2) We instantiate the metric on the current Internet topology of several countries, analyzing their network topologies with regards to Internet censorship. In the context of the influence to Internet reachability, we instantiate our metric using diverse groups of potentially influential countries (Section IV).
3) We perform extensive experiments using the contemporary Internet topology on both BGP, a waypoint network with intermediate nodes, and SCION, a path-aware Internet architectur, providing a comparative analysis for Internet censorship and global reachability (Section V).

ISBN 978-3-903176-63-8 © 2024 IFIP

II. Background

Internet Censorship. Internet censorship can be observed when an entity in power restricts its citizens or users from certain online communication or content, if it is deemed harmful, politically inappropriate, sensitive or legally noncompliant [1], [2]. Censorship could happen at various communication points: at the *end-point devices*, or *on-link*, by nodes that the traffic passes through [14]. Prior work indicates that the traversed nodes and the network topology do play a role for censorship [16], [18], [19], stressing the need for a quantitative evaluation of this influence. To that end, the scope and the focus of this article will be on the latter, *on-link* censorship.

Censorship Circumvention. In the face of widely deployed censorship techniques, a plethora of censorship circumvention methods have arisen [13]. Among others, one can use intermediate nodes in the network as a waypoint, to circumvent the censors' influence [14], for instance by utilizing a Virtual Private Network (VPN) connection.

Global Internet Reachability. Prior work shows a certain hegemonic influence on global Internet reachability [8], while many depend on Western countries to access common Internet destinations [10]. Not only can this jeopardize global reachability due to dependence on other countries and their Internet infrastructure [10], but it can also raise concerns about potential surveillance [20]–[22] and collateral damage [18].

SCION Next-generation Internet Architecture. SCION is a *next-generation Internet architecture* [12], [23], which is already deployed in production networks [24]. SCION groups ASes into Isolation Domains (ISDs), with shared governance institutions. An ISD provides a trust environment between the ASes in it, which could—among others—group around a common jurisdiction [12]. SCION is a multi-path and path-aware architecture [25]. In general, each node can have several end-to-end paths at its disposal, from which it can select any one of them. This contrasts with the BGP/IP Internet that is hop-by-hop routed.

III. Quantifying Avoidability in a Network

In this section, we introduce *Avoidability Potential*, drawing from classic graph theory metrics like group betweenness centrality.

A. Avoidability Potential

Avoidability Potential quantifies the potential of avoiding undesirable or potentially malicious nodes in the network. In this work, we focus on the topological organization of the Internet, providing an analysis at the inter-AS routing level.

Network Model. We model the network as a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$. Each node $v \in \mathcal{V}$ represents an Autonomous System (AS), whereas each edge $e \in \mathcal{E}$ represents a link between nodes. Edges between nodes are labeled with the standard business relationships on the Internet: *customer-provider* and *peer-peer* [26]. Furthermore, the nodes in the SCION network are grouped into ISDs. This does not affect the model of the network as a graph, but rather only provides grouping of nodes composing it.

Threat Model. In a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, communication between nodes $s, d \in \mathcal{V}$ can be censored, intercepted, blocked, or in any way tampered with. This interference can occur anywhere on their path due to unreliable, potentially malicious, simply or untrusted nodes. We assume that ASes pose a threat as a whole, with varying numbers of Byzantine ASes that may collude. While the motives and interests for Internet censorship are diverse, our model does not explicitly address them.

The Metric. We define set $\mathcal{S} \subset \mathcal{V}$ as the set of all *source* nodes from which paths of interest originate, and the set $\mathcal{D} \subset \mathcal{V}$ as the set of all *destinations*, where the paths terminate. Finally, we define $\mathcal{X} \subset \mathcal{V}$ as a set of ASes that should be avoided when communicating between nodes of interest.

Given the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, the set $\mathcal{X}$ of nodes whose avoidability is analyzed, a source node $s \in \mathcal{S}$ and a destination node $d \in \mathcal{D}$, we define $e_{\mathcal{X}}(s \to d)$ as a binary flag of whether a path between s and d exists, which completely circumvents nodes in $\mathcal{X}$. If it exists, we say that these nodes have the full potential of establishing a connection.

$$e_{\mathcal{X}}(s \to d) = \begin{cases} 1, & \exists r, \text{ a path } s \to d, \text{ s.t. } \forall x \in \mathcal{X}, x \notin r \\ 0, & \text{otherwise} \end{cases}$$

From there, we define the *Avoidability Potential* by allowing for all possible sources $s \in \mathcal{S}$ and destinations $d \in \mathcal{D}$. This yields the final metric, presented in the Equation (1).

$$AP_{\mathcal{X}}(\mathcal{S}, \mathcal{D}) = \frac{\sum\limits_{\substack{s \in \mathcal{S} \\ d \in \mathcal{D}}} e_X(s \to d)}{\|\mathcal{S}\| \cdot \|\mathcal{D}\|} \tag{1}$$

The value $\|S\| \cdot \|D\|$ in the Equation (1) is the number of all pairs of sources and destinations, which leads to a normalized value $AP_{\mathcal{X}}(\mathcal{S}, \mathcal{D}) \in [0, 1]$. Here, 1 means that the nodes from $\mathcal{D}$ can always receive traffic from the nodes in $\mathcal{S}$, without traversing any node in $\mathcal{X}$, whereas 0 would mean that this traffic would *always* traverse some of these nodes.

The above introduced metric is general, applicable to any graph and sets of nodes in the graph $\mathcal{S}$, $\mathcal{D}$ and $\mathcal{X}$. It is also independent of the network model, routing protocol, and captures architectures which allow for multiple paths. Thereupon we lay out two important applications of this metric, which are largely relevant for today's Internet: censorship resilience, and global Internet reachability.

B. Censorship Resilience Potential

We apply the *Avoidability Potential* to the case of censorship resilience, deriving the *Censorship Resilience Potential* metric, where the set $\mathcal{X}$ is the set of censoring nodes $\mathcal{C}$.

$$CRP_{\mathcal{C}}(\mathcal{S}, \mathcal{D}) = AP_{\mathcal{C}}(\mathcal{S}, \mathcal{D}) \tag{2}$$

Example: National Outflow Traffic. We can focus on a specific country by defining $\mathcal{C}$ as the set of ASes with interests or capabilities to censor outflow traffic. For outflow traffic originating within the country and heading to a foreign AS, all national ASes are sources, thus part of set $\mathcal{S}$. Similarly, ASes outside the country are in the destination set $\mathcal{D}$.

Towards a Metric Agnostic to Normative Claims. As mentioned briefly in Section II, censorship occurs globally for diverse reasons, including political motives. While one might consider that quantifying censorship must incorporate these interests and provide normative justifications, we present a metric that views all network nodes as potential censors without delving into individual motives or activities[1].

Defining Censoring ASes. In the case where censoring ASes are not known *a priori*, we define them based on their potential to choke the highest number of paths from $\mathcal{S}$ to $\mathcal{D}$. Following the intuition of the outflow traffic —an example relevant for the current Internet [15], [27]— we define the set of censors $\mathcal{C}$ as a subset of $\mathcal{S}$, which have the *highest potential* of choking outflow paths that go from $\mathcal{S}$ to $\mathcal{D}$. In other words, we focus on the situation where $\mathcal{S} \cap \mathcal{D} = \emptyset$ and $\mathcal{C} \subset \mathcal{S}$.

A border AS is an AS in $\mathcal{S}$, with at least one direct link to an AS outside of $\mathcal{S}$ [15]. Set $\mathcal{B}$ is the set of all border ASes.

$$\mathcal{B} = \{b \in \mathcal{S} | \exists e = (b, x) \in E \text{ s.t. } x \notin \mathcal{S}\} \tag{3}$$

Following the work by Leyba et al. [15], we adapt the concept of choke potential to capture the concept of path-awareness. For that, consider a subset of border ASes, $\mathcal{B}' \subset \mathcal{B}$. Their *Cumulative Choke Potential* (CPP) is the fraction of outflow paths that they could choke together. The rigorous definition of $CCP_{\mathcal{S},\mathcal{D}}(\mathcal{B}')$ is shown in Equation 4, and due to normalization yields to $CCP_{\mathcal{S},\mathcal{D}}(\mathcal{B}') \in [0, 1]$.

$$CCP_{\mathcal{S},\mathcal{D}}(\mathcal{B}') = \frac{\sum\limits_{\substack{s \in \mathcal{S} \\ d \in \mathcal{D}}} f_{\mathcal{B}'}(s \to d)}{\|\mathcal{S}\| \cdot \|\mathcal{D}\|} \tag{4}$$

$$f_{\mathcal{B}'}(s \to d) = \begin{cases} 1, & \forall r, \text{ a path } s \to d, \exists b' \in \mathcal{B}', \textit{ s.t. } b' \in r \\ 0, & \text{otherwise} \end{cases}$$

We define the set of censoring ASes $\mathcal{C}$ as a subset of $\mathcal{B}$ with a cardinality of $\|\mathcal{C}\| = N$, collectively capable of choking the highest number of outflow paths. As an intuition, if all border ASes were to censor, their cumulative choke potential would be 1, resulting in $CRP_{\mathcal{C}}(\mathcal{S}, \mathcal{D}) = 0$. However, strict enforcement of censorship across all border ASes is challenging. Thus, we conduct experiments across various countries with different values of N, elaborated on in Section IV.

Algorithm for *CRP* Metric. In a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, one must define the sets of source and destination nodes, $\mathcal{S} \subset \mathcal{V}$ and $\mathcal{D} \subset \mathcal{V}$, respectively. These sets can be chosen arbitrarily or based on node properties, such as country of origin. The set of censoring ASes $\mathcal{C} \subset \mathcal{V}$ can be defined in two ways.

In the first method, $\mathcal{C}$ is known *a priori*. The metric's value is determined by calculating the fraction of paths from $\mathcal{S}$ to $\mathcal{D}$ that avoid censoring ASes. The algorithm pipeline for this method is depicted in Figure 1, with graph and set inputs marked in gray, intermediate steps in purple, and the outputs in yellow.

[1]We do however note that our metric can also be applied to a set of censoring ASes that are *a priori* labeled as such.

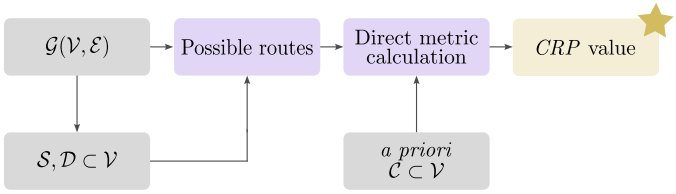

Fig. 1: *Censorship Resilience Potential* (CRP): the algorithm pipeline, with censoring ASes known *a priori*.

The second method determines the value of CRP using *Cumulative Choke Potential* (CCP). In this case, the underlying assumption is that censoring ASes would be border ASes from the set $\mathcal{S}$. The CCP value of the subset of them provides us with both the set of censoring ASes $\mathcal{C}$ with high potential of cumulatively choking outflow traffic, and the value of the final metric. The full algorithm pipeline of this method is laid out on Figure 2, with the same color-coding from Figure 1.

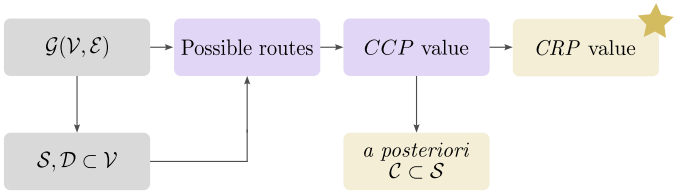

Fig. 2: *Censorship Resilience Potential* (CRP): the algorithm pipeline, where the *Cumulative Choke Potential* (CCP) is used for defining the set of censoring ASes $\mathcal{C}$ *a posteriori*.

***CRP* as Means for Comparative Analysis.** Our metric aims to facilitate comparative analysis of different network models and Internet architectures. If the set of censoring ASes is known beforehand, the method outlined in Figure 1 should be applied universally. However, if the set $\mathcal{C}$ is not predetermined, it should be defined independently. For a comprehensive analysis, $\mathcal{C}$ should be defined according to the pipeline depicted in Figure 2 for each architecture. We utilize this approach in our simulation, discussed further in Section IV, yielding results suitable for comparing BGP, waypoint models, and SCION.

C. Global Reachability Potential

Our second goal is to gauge the potential of reaching the global Internet, while avoiding undesirable nodes in the network. We achieve this independently of the specific network topology, or the Internet architecture, by adopting the *Avoidability Potential* metric to this use case.

The Metric. Given the graph representing the global Internet, $\mathcal{G}(\mathcal{V}, \mathcal{E})$, it is possible that certain nodes are more central to for global connectivity than others. To measure how much influence a group of nodes $\mathcal{X} \subset \mathcal{V}$ has on nodes $\mathcal{S} = \mathcal{V} \setminus \mathcal{X}$ to establish paths with each other, we employ the *Avoidability Potential* metric, for convenience calling it *Global Reachability Potential*.

$$GRP_{\mathcal{X}}(\mathcal{S}, \mathcal{S}) = AP_{\mathcal{X}}(\mathcal{S}, \mathcal{S}), \ \mathcal{S} = \mathcal{V} \setminus \mathcal{X} \tag{5}$$

Example: Collateral Damage of Internet Censorship. While censorship techniques primarily target specific network nodes [7], they can also lead to collateral damage affecting other nodes beyond the intended scope [28]. For instance, Acharya et al. suggest that countries known for censorship may impact global reachability [18]. Our metric can analyze this collateral censorship damage at the AS level, extending its applicability to such scenarios.

Example: Influence of Hegemonic Groups. Several authors note that a small number of ASes serve as global transit networks [29], raising concerns about their hegemonic influence on global Internet reachability [8]. The *Global Reachability Potential* metric can analyze the potential for circumventing such influential nodes, offering comparative analysis of network models and quantitative evidence of Internet routing centralization or democratization.

IV. Experimental Evaluation

In this section, we outline the setup and approach for extensive experiments, whose results we elaborate on in Section V.

Overview of Analyzed Network Models. In our study, we apply our metrics to three different network models. First, we look at the current BGP/IP Internet design, and without considering any routing attacks. Second, we consider a scenario where waypoint mechanisms are widely used to bypass undesirable nodes in the network. Lastly, we examine SCION, a path-aware next-generation Internet architecture, where each end-host can choose the whole end-to-end path.

A. Datasets

For all our experiments, we use the datasets that represents real and contemporary relationships between ASes.

AS Relationships. We utilize the CAIDA AS Relationships dataset [30] to model the network graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$. This dataset offers the Internet's topology, employed directly in all BGP and waypoint model simulations. We maintain consistency in our SCION simulation by employing the same topology, ensuring comparable and relevant results.

AS Country Origin. To accurately determine the country origin of each AS, we employ two datasets. First, the CAIDA AS to Organizations Mappings offers legal entity country origins for ASes [31]. Second, the RIPEStat Geo Map dataset [32] provides physical locations where ASes announce BGP prefixes. We account for Tier-1 ASes being present in multiple countries, interconnected through branches.

Waypoints in the Network. For the waypoint network model, we utilize the anonymous dataset from MaxMind. This dataset identifies ASes previously associated with potential host anonymization services, such as VPN or Tor nodes [33].

B. Country Network

We apply the *Censorship Resilience Potential* metric to network nodes according to their country of origin. Our selection of countries encompasses diversity based on various indicators, such as geographical location, population size, national network size, and the *Internet Freedom Score* (IFS) [34].

Nodes Forming a Country Network. Let $\mathfrak{X}$ denote a country of interest, and $\mathcal{K}$ represent the set of ASes originating from $\mathfrak{X}$. We define $\mathcal{K}$ as follows:

$$\mathcal{K} = \{k \in \mathcal{V} |\ \text{country}(k) = \mathfrak{X}\} \tag{6}$$

We use set $\mathcal{K}$ to define a country's network, excluding outliers. Inspired by the study of Guillermo et al. on the global Internet [35], we identify *islands* in a country network, labeling the largest connected component as the country network.

C. SCION Topology

Although SCION is already deployed, its current production network footprint does not yet reach the scale of the BGP infrastructure. To address this, we construct the SCION topology using the CAIDA AS Relationship dataset [30], ensuring it reflects real-world deployment and remains comparable to the BGP/IP network.

Core ASes. We define core ASes as per the analysis by Krähenbühl et al. [24], based on the customer cone size of ASes. When determining the value of the *Global Reachability Potential*, we use the graph of core ASes, as discussed in the remainder of this paper.

Grouping into ISDs. Grouping nodes into an ISD is essential for applying the *Censorship Resilience Potential* metric, especially on a per-country basis. We assume that ASes connected to a country's network infrastructure naturally form a connected component, creating a "national" ISD—or a group of ISDs in the general case [12]. We keep links between all ASes in an ISD and disregard links to ASes outside the ISD.

D. Simulation on Diverse Network Models

In this section we comment on the implementation details.

Censoring ASes. We conduct inter-AS BGP simulation using routing tree algorithm by Gill et al. to determine preferred paths between ASes in the Internet topology [36]. In the BGP network model, we define $\mathcal{C} = \mathcal{C}_{\mathcal{BGP}}$ as a subset of border ASes with the highest potential to restrict outflow traffic. Drawing on Leyba et al.'s work, we attribute potentially choked paths to the last border ASes [15]. Additionally, since the waypoint model shares the same topology and routing algorithm as the BGP model, we employ the same set of censoring ASes for consistency. Finally, for the SCION network topology, we select a subset of border ASes with the highest customer cone size, forming the set of censoring nodes, denoted as $\mathcal{C} = \mathcal{C}_{\mathcal{SCION}}$. This method is an effective heuristic for selecting nodes with the highest potential to cumulatively choke outflow paths, as it considers the customer cone size of each border AS in the country network.

Censorship Resilience Potential. Once the set of censoring ASes $\mathcal{C}$ is established, we compute the *Censorship Resilience Potential* metric for both the BGP and waypoint models by assessing the fraction of paths not intercepted by nodes in $\mathcal{C}$. In SCION, sources have the freedom to choose the entire end-to-end path. Thus, we determine whether exists a path that can leave the ISD while avoiding nodes in $\mathcal{C}$.

	Number of ASes	Border ASes (BGP/Waypoint)	Border ASes (SCION)
Brazil	8174	2285 (28%)	243 (3%)
China	534	94 (18%)	21 (4%)
India	2537	209 (8%)	36 (1%)
Iran	481	25 (5%)	5 (1%)
Russia	4957	1139 (23%)	82 (2%)
Switzerland	654	308 (47%)	28 (4%)
U.K.	1562	861 (55%)	62 (4%)
United States	17934	2173 (12%)	236 (1%)

TABLE I: Country network statistics: number of total and border ASes in a country network, across all analyzed models.

Global Reachability Potential. Once the set of nodes analyzed for global influence $\mathcal{X}$ is defined, we compute the *Global Reachability Potential* metric for both the BGP and waypoint models by assessing the fraction of paths not intercepted by nodes in $\mathcal{X}$. For SCION, it is enough analyzing the interconnectedness of the core ASes, as they are crucial for the global Internet reachability.

V. Experimental Results

We apply our metric to two distinct cases: censorship resilience of various countries, and global Internet reachability, commenting on their results in this section. For extended results refer to Section Availability.

A. National Censorship Resilience Potential

We assess the *Censorship Resilience Potential* metric across BGP, waypoint model, and SCION, incorporating diverse countries. Our findings underscore the critical influence of network topology on Internet censorship. Figure 3 presents our results, which we analyze further in this section.

Border ASes as Central Choking Points. Our findings reveal that even a small number of a country's border ASes can significantly restrict outflow paths in the current Internet. With as few as 20 border ASes, up to 50% of outflow paths can be choked, irrespective of the country network's size. For instance, in BGP, the United States could potentially choke 26% of outflow paths ($CRP = 0.74$) with just one AS. We complement this analysis with network statistics in Table I, indicating network size and the number of border ASes across all models. For example, comparing the Iranian and Swiss BGP-network topologies, we observe Iran's more centralized model with fewer border ASes, corroborating its centralized censorship model [3], [17]. Conversely, Switzerland exhibits a higher density of border ASes. This comparison offers insights into other countries' censorship efforts and the potential collaboration among ASes in such endeavors.

Network Topology Matters. The results confirm that the number of border ASes is not the sole influential factor. The interconnectedness of ASes within the network also matters, regardless of the Internet architecture. In other words, even with numerous exit points from a country network, the routing among them might not be evenly distributed, leading to dependency on a small number of ASes. For instance, the United States, with nearly 18'000 ASes, exhibits a relatively low CRP value of 0.32, attributed to only 5 ASes and their routing influence.

	BGP	Waypoint	SCION
United States	0.59	0.92	0.9951
Five Eyes	0.52	0.88	0.9941
European Union	0.87	0.98	0.9975
Iran, China, Russia	0.98	0.99	0.9995

TABLE II: Results of the *Global Reachability Potential*, with various groups of countries analyzed for global influence.

Path-selection and Censorship Resilience. Multiple exit paths from a country offer potential for enhancing censorship resilience, yet they are often underutilized. In path-aware technologies like the waypoint model, the likelihood of a single AS controlling all outflow paths is noticeably reduced compared to BGP. In SCION, the lower number of Border ASes, due to technical and governance factors, results in fewer exit points in the network (see Table I). However, our analysis demonstrates that path-selection mechanisms can significantly mitigate the impact of censoring Border ASes by offering alternatives to bypass undesired ASes.

B. Global Internet Reachability

We assess the *Global Reachability Potential* metric outlined in section III-C across BGP, waypoint model, and SCION. The selected influential country groups align with previous studies [8], [29], identifying nations with either negative impact on Internet reachability or potential for collateral damage due to censorship. Results are summarized in Table II.

Nodes Centrality. The results indicate that ASes from the United States and the Five Eyes countries serve as transit nodes for over 40% of all paths between nodes in different countries. Similarly, ASes from the European Union play a central role in global reachability. We also assess the *Global Reachability Potential* of Iran, China, and Russia, commonly discussed in the context of censorship [3], [4], [37] and its potential collateral damage [18]. Our analysis reveals that they do not exert significant influence on global reachability across the three network models analyzed, as they lack central positioning in the current Internet topology.

Path-selection and Global Reachability. Results in Table II reveal variations among the three analyzed models. The impact of examined groups on global Internet reachability is lower in the waypoint model compared to BGP. Multiple waypoint hosts in this model suggest the existence of various paths, increasing the potential to bypass undesired ones. However, as several waypoint nodes originate from the United States, achieving full *Global Reachability Potential* remains challenging. Additionally, SCION's end-to-end path-awareness offers means to circumvent undesirable nodes, resulting in high *Global Reachability Potential* for all analyzed country groups.

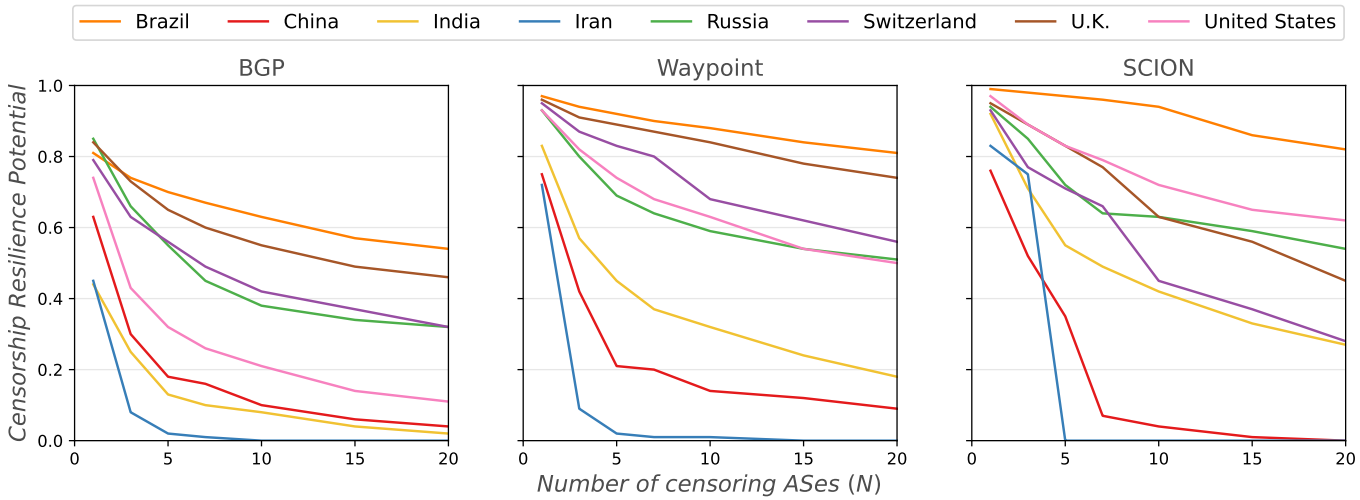

Fig. 3: *Censorship Resilience Potential* for BGP, waypoint model, and SCION, presented for various countries and varying number of censoring ASes, N.

VI. Related Work

Country Network Analysis. In the realm of Internet censorship, prior studies by various authors [38] emphasize the significance of network topology. For instance, Ensafi et al. note that Tor traffic, often censored in China, bypasses censorship when entering via CERNET [37]. Gill et al. characterize Iran's network as centralized [7], while Salamatian et al. highlight the limited direct links of Iran's network to foreign ASes, suggesting strategic use of BGP for censorship [17]. Additionally, Wählisch et al. employ a sector-based approach to assess betweenness centrality in the German network [39].

Control of National Outflow Traffic. Roberts et al. developed a measure of network complexity, unveiling underlying properties indicative of a country's censorship capabilities [16]. Similarly, Leyba et al. found that the number of nodes capable of choking a significant fraction of outflow paths is not only low but also decreasing over time [15].

Global Internet Reachability. Other researchers examined global Internet reachability by assessing betweenness centrality on a global scale, pinpointing nodes and countries pivotal for global connectivity [8], [29], and delving into the potential collateral damage from censorship efforts [18].

Broadened Prior Work and Contributions. Our metric builds on prior work to create a comprehensive tool for quantifying censorship resilience and global Internet reachability. It is versatile across various network models and topologies, including path-aware Internet architectures. Importantly, it does not require predefined censoring ASes for assessing censorship resilience.

Next-generation Internet Architectures and Censorship. We underscore the significance of scrutinizing next-generation Internet architectures in the context of censorship and Internet reachability. While Kohler [40] and Wrana et al. [19] explored this qualitatively, our main contribution is a quantitative metric suitable for comparative analysis.

VII. Discussion

Routing Attacks on BGP. When assessing *Censorship Resilience Potential* and *Global Reachability Potential* metrics on BGP and the waypoint model, we establish legitimate paths between any two nodes, without considering routing attacks by malicious actors. Such attacks could involve redirecting traffic [41], jeopardizing both censorship resilience and global Internet accessibility.

Waypoint Model on the Internet. Our waypoint model sheds light on the influence of systems like VPN connections in bypassing undesirable nodes. However, it might oversimplify censorship circumvention by suggesting reliance on waypoint service providers, whereas censoring entities often block IP addresses from such providers. Nevertheless, it offers quantitative evidence of the advantages of multiple paths in Internet routing.

Internet Deployments. BGP serves as the exclusive inter-domain routing protocol, rendering our BGP results directly relevant to the current Internet. Conversely, SCION's limited deployment [24] positions our findings as a future prospect. Moreover, waypoints, such as VPN connections, can complement SCION since their deployment is independent and compatible.

Policy Impacts. We avoid making normative claims about Internet censorship or global reachability. Instead, we provide an objective metric for their evaluation, serving as a quantitative tool. This approach can offer policymakers insights into network technology design, development, and deployment.

VIII. Conclusion

In this paper, we have demonstrated how network topology and Internet architecture can affect a country's resilience to Internet censorship and its global reachability, highlighting dependencies on specific network nodes. We have proposed a novel approach that utilizes quantitative network metrics to

evaluate these aspects of today's Internet. We evaluated the metric across diverse Internet models using Border Gateway Protocol (BGP), a model of waypoints based on Virtual Private Networks (VPNs), and SCION path-aware architecture. Our results underscore the importance of network topology and suggest that path-aware architectures could democratize global routing, potentially enhancing censorship resilience.

Ethics Statement

In this work we use datasets already publicly available, and do not conduct any additional Internet measurements that may raise ethical concerns.

Availability

The accompanying code repository and extended results are available at: https://github.com/IvanovicM/charting-censorship

References

[1] J. L. Hall, M. D. Aaron, A. Andersdotter, B. Jones, N. Feamster, and M. Knodel, "A Survey of Worldwide Censorship Techniques," Internet Engineering Task Force, Internet-Draft draft-irtf-pearg-censorship-09, Jan. 2023, work in Progress.

[2] C. S. Leberknight, M. Chiang, H. V. Poor, and F. Wong, "A taxonomy of internet censorship and anticensorship," *Fifth International Conference on Fun with Algorithms*, 2010.

[3] S. Aryan, H. Aryan, and J. A. Halderman, "Internet censorship in iran: A first look," in *3rd USENIX Workshop on Free and Open Communications on the Internet (FOCI 13)*. Washington, D.C.: USENIX Association, Aug. 2013.

[4] R. Ramesh, R. S. Raman, M. Bernhard, V. Ongkowijaya, L. Evdokimov, A. Edmundson, S. Sprecher, M. Ikram, and R. Ensafi, "Decentralized control: A case study of Russia," in *Network and Distributed System Security*. The Internet Society, 2020.

[5] K. Singh, G. Grover, and V. Bansal, "How India censors the web," in *Web Science*. ACM, 2020.

[6] B. Marczak, N. Weaver, J. Dalek, R. Ensafi, D. Fifield, S. McKune, A. Rey, J. Scott-Railton, R. Deibert, and V. Paxson, "An analysis of China's "Great Cannon"," in *Free and Open Communications on the Internet*. USENIX, 2015.

[7] P. Gill, M. Crete-Nishihata, J. Dalek, S. Goldberg, A. Senft, and G. Wiseman, "Characterizing web censorship worldwide: Another look at the OpenNet Initiative data," *Transactions on the Web*, vol. 9, no. 1, 2015.

[8] J. Karlin, S. Forrest, and J. Rexford, "Nation-state routing: Censorship, wiretapping, and BGP," *arXiv*, 2009.

[9] A. Shah, R. Fontugne, and C. Papadopoulos, "Towards characterizing international routing detours," in *Proceedings of the 12th Asian Internet Engineering Conference*, ser. AINTEC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 17–24.

[10] A. Edmundson, R. Ensafi, N. Feamster, and J. Rexford, "Nation-state hegemony in internet routing," in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, ser. COMPASS '18. New York, NY, USA: Association for Computing Machinery, 2018.

[11] Y. Rekhter, S. Hares, and T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 4271, Jan. 2006.

[12] L. Chuat, M. Legner, D. A. Basin, D. Hausheer, S. Hitz, P. Müller, and A. Perrig, *The Complete Guide to SCION - From Design Principles to Formal Verification*, ser. Information Security and Cryptography. Springer, 2022. [Online]. Available: https://doi.org/10.1007/978-3-031-05288-0

[13] M. C. Tschantz, S. Afroz, Anonymous, and V. Paxson, "SoK: Towards grounding censorship circumvention in empiricism," in *Symposium on Security & Privacy*. IEEE, 2016.

[14] S. Khattak, T. Elahi, L. Simon, C. M. Swanson, S. J. Murdoch, and I. Goldberg, "SoK: Making sense of censorship resistance systems," *Privacy Enhancing Technologies*, vol. 2016, no. 4, pp. 37–61, 2016.

[15] K. G. Leyba, B. Edwards, C. Freeman, J. R. Crandall, and S. Forrest, "Borders and gateways: Measuring and analyzing national as chokepoints," in *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies*, ser. COMPASS '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 184–194.

[16] H. Roberts, D. Larochelle, R. Faris, and J. Palfrey, "Mapping local Internet control," in *Computer Communications Workshop*. IEEE, 2011.

[17] L. Salamatian, F. Douzet, K. Salamatian, and K. Limonier, "The geopolitics behind the routes data travel: a case study of Iran," *Journal of Cybersecurity*, vol. 7, no. 1, p. tyab018, 08 2021.

[18] H. B. Acharya, S. Chakravarty, and D. Gosain, "Few throats to choke: On the current structure of the internet," in *2017 IEEE 42nd Conference on Local Computer Networks (LCN)*, 2017, pp. 339–346.

[19] M. Wrana, D. Barradas, and N. Asokan, "The spectre of surveillance and censorship in future internet architectures," *arXiv*, 2024.

[20] A. Edmundson, R. Ensafi, N. Feamster, and J. Rexford, "A first look into transnational routing detours," in *Proceedings of the 2016 ACM SIGCOMM Conference*, ser. SIGCOMM '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 567–568.

[21] ——, "Characterizing and avoiding routing detours through surveillance states," *CoRR*, vol. abs/1605.07685, 2016. [Online]. Available: http://arxiv.org/abs/1605.07685

[22] J. Obar and A. Clement, "Internet Surveillance and Boomerang Routing: A Call for Canadian Network Sovereignty," *SSRN Electronic Journal*, 2013.

[23] S. Bechtold and A. Perrig, "Accountability in future internet architectures," *Commun. ACM*, vol. 57, no. 9, p. 21–23, sep 2014.

[24] C. Krähenbühl, S. Tabaeiaghdaei, C. Gloor, J. Kwon, A. Perrig, D. Hausheer, and D. Roos, "Deployment and scalability of an inter-domain multi-path routing infrastructure," in *Proceedings of the 17th International Conference on Emerging Networking EXperiments and Technologies*, ser. CoNEXT '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 126–140.

[25] B. Trammell, J.-P. Smith, and A. Perrig, "Adding path awareness to the internet architecture," *IEEE Internet Computing*, vol. 22, no. 2, pp. 96–102, 2018.

[26] L. Gao and J. Rexford, "Stable internet routing without global coordination," *IEEE/ACM Transactions on Networking*, vol. 9, no. 6, pp. 681–692, 2001.

[27] X. Xu, Z. M. Mao, and J. A. Halderman, "Internet censorship in China: Where does the filtering occur?" in *Passive and Active Measurement Conference*. Springer, 2011, pp. 133–142.

[28] Sparks, Neo, Tank, Smith, and Dozer, "The collateral damage of Internet censorship by DNS injection," *SIGCOMM Computer Communication Review*, vol. 42, no. 3, pp. 21–27, 2012.

[29] R. Fontugne, A. Shah, and E. Aben, "The (thin) bridges of AS connectivity: Measuring dependency using AS hegemony," *CoRR*, vol. abs/1711.02805, 2017. [Online]. Available: http://arxiv.org/abs/1711.02805

[30] Center for Applied Internet Data Analysis, "AS Relationships (serial-2)."

[31] ——, "Inferred AS to Organization Mapping Dataset."

[32] RIPEstat, "RIPE Stat."

[33] MaxMind, "GeoIP2 Anonymous IP Database."

[34] Freedom House, "Internet Freedom Scores."

[35] G. Baltra and J. Heidemann, "What is the internet? (considering partial connectivity)," University of Southern California, Tech. Rep., 2022.

[36] P. Gill, M. Schapira, and S. Goldberg, "Modeling on quicksand: Dealing with the scarcity of ground truth in interdomain routing data," *SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 1, p. 40–46, 2012.

[37] R. Ensafi, P. Winter, A. Mueen, and J. R. Crandall, "Analyzing the Great Firewall of China over space and time," *Privacy Enhancing Technologies*, vol. 2015, no. 1, 2015.

[38] A. Master and C. Garman, "A worldwide view of nation-state Internet censorship," in *Free and Open Communications on the Internet*, 2023. [Online]. Available: https://www.petsymposium.org/foci/2023/foci-2023-0008.pdf

[39] M. Wählisch, T. Schmidt, M. de Brün, and T. Häberlen, "Exposing a nation-centric view on the german internet – a change in perspective on the as level," in *International Conference on Passive and Active Network Measurement*, vol. 7192, 03 2012.

[40] K. Kohler, "One, Two, or Two Hundred Internets? The Politics of Future Internet Architectures," *CSS Cyberdefense Reports*, 2022.

[41] S. L. Murphy, "BGP Security Vulnerabilities Analysis," RFC 4272, Jan. 2006.