Destination Unreachable: Characterizing Internet Outages and Shutdowns

Zachary S. Bischof
Georgia Tech

Kennedy Pitcher
UC San Diego

Esteban Carisimo
Northwestern University

Amanda Meng
Georgia Tech

Rafael Bezerra Nunes
Yale University

Ramakrishna Padmanabhan[*]
Amazon Web Services

Margaret E. Roberts
UC San Diego

Alex C. Snoeren
UC San Diego

Alberto Dainotti
Georgia Tech

ABSTRACT

In this paper, we provide the first comprehensive longitudinal analysis of government-ordered Internet shutdowns and spontaneous outages (i.e., disruptions not ordered by the government). We describe the available tools, data sources and methods to identify and analyze Internet shutdowns. We then merge manually curated datasets on known government-ordered shutdowns and large-scale Internet outages, further augmenting them with data on real-world events, macroeconomic and sociopolitical indicators, and network operator statistics. Our analysis confirms previous findings on the economic and political profiles of countries with government-ordered shutdowns. Extending this analysis, we find that countries with national-scale spontaneous outages often have profiles similar to countries with shutdowns, differing from countries that experience neither. However, we find that government-ordered shutdowns are many more times likely to occur on days of mobilization, coinciding with elections, protests, and coups. Our study also characterizes the temporal characteristics of Internet shutdowns and finds that they differ significantly in terms of duration, recurrence interval, and start times when compared to spontaneous outages.

CCS CONCEPTS

• **Networks** → **Public Internet**; *Network monitoring*;

KEYWORDS

Internet shutdowns, Internet outages, Internet reliability, Internet censorship

ACM Reference Format:
Zachary S. Bischof, Kennedy Pitcher, Esteban Carisimo, Amanda Meng, Rafael Bezerra Nunes, Ramakrishna Padmanabhan, Margaret E. Roberts, Alex C. Snoeren, and Alberto Dainotti. 2023. Destination Unreachable: Characterizing Internet Outages and Shutdowns. In *ACM SIGCOMM 2023 Conference (ACM SIGCOMM '23), September 10–14, 2023, New York, NY, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3603269.3604883

1 INTRODUCTION

The Internet research community has made significant efforts to improve the performance and reliability of the Internet's infrastructure and its protocols. However, these efforts lose their relevance when the Internet's entire technology stack is taken away, sometimes for days at a time. Unfortunately, this is the harsh reality for billions of individuals around the globe in countries where governments intentionally shut off access to the Internet [39]. While there is a long history of discussions related to the right to Internet access, its significance increased following the Internet shutdowns in 2011 in Egypt, Libya, and Syria during the Arab Spring. Despite the UN's resolution in 2016 condemning governments intentionally shutting down access to the Internet [55], shutdowns continue to be prevalent and increasingly frequent [27, 44]. Though the research community is aware of this phenomenon, research on the topic often focuses on a single case study or series of related events occurring in one country. As a result, both practitioners and academics lack a systematic understanding of the characteristics of Internet shutdowns as well as the tools—and their limitations—available to analyze and document them.

This work aims to improve the field's understanding of national, large-scale Internet *shutdowns* (i.e., intentionally ordered by governments) as a whole and how they differ in comparison to *spontaneous outages* (i.e., connectivity disruptions that are not government-ordered) of similar scale. To achieve this goal, we perform the first systematic analysis of Internet shutdowns by studying them alongside spontaneous outages. We provide a novel dataset that combines (1) Internet measurement data from the Internet Outage Detection and Analysis (IODA) research project [30], (2) Internet shutdown data from the digital rights organization Access Now [2], (3) political and socioeconomic data, and (4) data on network operator statistics. This dataset of national-scale events spans 155 countries over 4 years and includes 219 Internet shutdowns and 714 spontaneous outages.

Similar to previous works that demonstrate a link between Internet censorship and authoritarianism [7], we find that Internet shutdowns tend to occur in countries that are more authoritarian. Extending this analysis to include spontaneous outages, we find that the economic and political profiles of countries with spontaneous outages are in some ways similar to those of countries that experience shutdowns, often differing from countries that experience neither. We find that instead what seems to distinguish between shutdowns and spontaneous outages are moments of mobilization,

[*] Work done while at UC San Diego.

ACM SIGCOMM '23, September 10–14, 2023, New York, NY, USA

ACM ISBN 979-8-4007-0236-5/23/09.
https://doi.org/10.1145/3603269.3604883

the infrastructure of the Internet itself, and temporal and technical fingerprints that suggest intention. Consistent with a larger literature on political censorship [23], we find that shutdowns are much more likely to occur on days of transition and turmoil, including during elections, protest events, and coups. Furthermore, these days are not more likely to have a spontaneous outage. Control of the address space is an additional important predictor of shutdowns. Finally, indicators collected from telemetry data suggest human intervention, including the timing of shutdowns (on the hour, systematic recurrence, and increased frequencies on weekdays) tend to be associated with shutdowns but not spontaneous outages.

Our key contributions are: *(i)* The first longitudinal interdisciplinary study combining research from Internet measurement and political science to provide insight into the nature, origin, and identifying signatures of politically-motivated Internet shutdowns. Put into practice, the insights we provide through this systematic research can aid governments, intergovernmental organization, policy makers, and practitioners in tracking, reporting, and litigating on Internet shutdowns, which violate digital human rights. By comparing shutdowns and spontaneous outages, our analysis provides key characteristics for Internet freedom organizations to look for when working to efficiently and rapidly identify Internet shutdowns. Additionally, this foundational work can inform future studies on rapid identification of shutdowns using predictive models. *(ii)* We compile and share an unprecedentedly detailed and comprehensive dataset of Internet shutdowns and spontaneous outages spanning four years[1]. We hope these data will enable further study of this topic. We intend to continue populating our dataset and open it for contributions and annotations from other sources.

2 BACKGROUND AND RELATED WORK

The term "Internet shutdown" can be used to describe a variety of government-ordered Internet restrictions. Governments have sometimes mandated the complete disconnection of users from the Internet [19, 47], while at other times, they have blocked specific websites [57], social media platforms [6], and circumvention tools [48]. Our analysis of shutdowns (and use of this term in this paper) focuses specifically on instances where users are completely disconnected from the Internet.

Government-mandated Internet restrictions have been studied and measured from a wide range of angles. The extensive measurements and reports of incidents around the world provide researchers with a multitude of datasets and case studies to explore Internet shutdowns. In addition to IODA [30], there are several research groups from academia as well as the private and public sectors that measure Internet blocking and censorship. Examples include Censored Planet [12], the Open Observatory for Network Interference (OONI) [45], Cloudflare Radar [15], Google Jigsaw and the Google Transparency Report projects [25, 26], and Mozilla telemetry data [40]. Censored Planet measures network interference on the TCP/IP, DNS, and HTTP(S) protocols using remote measurement techniques. OONI collects data from users around the world who download their OONI probe application and run tests to detect the blocking of websites, applications, and services. Cloudflare Radar uses data from Cloudflare's network and public DNS resolver to identify trends and network outages. Google Transparency Report provides aggregated data of traffic to Google products (e.g., Google Search, Gmail, YouTube) at the country level. Mozilla telemetry data provides Firefox browser usage data at the country and city level. These measurement groups provide data that is complementary to IODA's monitoring of the connectivity of Internet infrastructure at the country, region, and operator level.

There are several studies investigating the censorship methods used by governments, from full disconnection of users (which we refer to here as shutdowns) to application-specific bans. The enforcement of application-layer bans have achieved notoriety in recent years. Iran was found to apply multiple application-layer restrictions where Aryan *et al.* [4] discovered host–based blocking, keyword filtering, DNS hijacking, and service-specific throttling of sites in the Alexa Top-500 websites. In Kazakhstan, where the state *de facto* controls the Internet via state-owned providers, Raman *et al.* [52] found that the government forces users to install a custom root certificate to be able to apply fine-grained content blocking. Pearce *et al.* [49] developed Iris as a method to detect manipulation of DNS queries that could reveal censorship. Zittrain *et al.* [59] studied state-sponsored Internet censorship and discovered that governments often ban content of foreign states.

While our work is, to the best of our knowledge, the first longitudinal characterization of political shutdowns and spontaneous outages, previous studies investigated isolated shutdowns or a series of related shutdowns. Dainotti *et al.* [18] identified country-wide network outages as a government initiative to suffocate protests during the Arab spring. More recently, Padmanabhan *et al.* [48] showed the enforcement of multiple Internet bans, including network disconnections during the 2021 coup in Myanmar. Eneyew Ayalew [5] examined three years of government-mandated shutdowns in Ethiopia. In the social sciences, Freyburg and Garbe [23] detected a correlation in sub-Saharan Africa between state-owned Internet providers and government-mandated shutdowns. Howard *et al.* [28] documented an analysis of a hand-curated dataset of political censorship events in countries from 1995-2011, including both website blocking and outages. Gohdes [24] analyzed the political determinants of Internet shutdowns in Syria. This body of related work provides important related censorship measurements and case studies on methods of implementing a government shutdown and foreground specific cases of shutdowns in the Middle East, Africa, and Asia. In this paper, we focus on studying full network shutdowns as opposed to events that only involve throttling or the blocking of specific websites or applications. In contrast to the case-study literature, we look at longitudinal data spanning across countries to identify cross-cutting characteristics of Internet shutdowns.

3 DATASETS

The focus of this paper is to better understand the conditions under which large-scale (i.e., nation-scale) full network shutdowns occur and how these conditions compare to countries that experience spontaneous outages of a similar scale as well as those without large-scale service disruptions. Accomplishing this task requires combing various sources of indicators spanning disciplinary domains. In this

[1] The dataset and code used in our analyses are available at https://github.com/InetIntel/internet_outages.

section, we describe the tools we use to record Internet disruptions as well as the datasets used in our analysis.

3.1 IODA

The Internet Outage Detection and Analysis (IODA) project [30] became operational in 2016. IODA publishes, in near-realtime, data on multiple indicators of Internet connectivity. Visitors to the IODA website can use the publicly available API, dashboards and graphs to investigate potential outages for particular countries, subnational regions (e.g., state or province), and Autonomous Systems (ASes).

IODA extracts from measurements and publishes three types[2] of time series signals: *(i)* Border Gateway Protocol (BGP), *(ii)* Active Probing, and *(iii)* Telescope. For each of these, IODA generates automated alerts when it detects a drop compared to the median of a historical sliding window. The threshold for what is considered a drop is specific to each signal, as some (e.g., BGP) are more stable than others (e.g., Telescope). Though these alerts are relatively simplistic and have false-positives, they serve as a useful starting point for investigating via manual inspection the countries, regions, and ASes that are potentially undergoing an outage. Drops affecting a given country (or region or AS) that are visible across multiple IODA signals that are overlapping in time, are good indicators of a network outage. An IODA signal can experience a drop due to a non-network-outage event (i.e., a false positive), but due to the signals' complementary nature, it is highly unlikely that more than one signal would experience simultaneous false-positive drops.

3.1.1 Data sources. The following paragraphs describe IODA's signals as well as how their corresponding alerts are generated. Additional details can be found on IODA's Help page [35].

BGP. For its BGP signal, IODA analyzes data from all RouteViews [43] and RIPE RIS collectors [53] using BGPStream [46] with BGPView [9]. For each time bin, IODA calculates the total number of "full-feed" peers that observe each routable prefix. A peer is considered full-feed if it has more than 400k IPv4 prefixes and/or more than 10k IPv6 prefixes. A prefix is considered visible if it is observed by at least 50% of the full-feed peers. IODA uses this data to calculate the total number of visible /24s per country, region, and AS every 5 minutes. IODA generates an automatic BGP alert when the number of /24 blocks that are visible to at least 50% of full-feed peers drops below 99% of the median of the previous 24 hour time-window.

Active Probing. IODA conducts active measurements using a technique similar to Trinocular [51], probing approximately 4.2M /24 blocks at least once every 10 minutes via ICMP packets. Using the Trinocular measurement and inference technique, IODA labels each /24 block as up, down, or unknown. After each 10-minute cycle, IODA calculates the number of /24s that are considered active for each country, subnational region, and AS. IODA generates an Active Probing alert when the current number of /24 blocks considered active drops below 80% of the median of the preceding 7 day time-window.

Telescope. To obtain the Telescope signal, IODA analyzes traffic received by a network telescope. IODA applies multiple anti-spoofing heuristics [18] and noise reduction filters to the raw traffic to create a set of valid packets. For each valid packet, IODA uses geolocation databases and AS lookups to map a packet's source IP address to a geographic location and AS. For each country, region, and AS view, the IODA dashboard displays the number of unique source IP addresses observed in each 5 minute bin. Though IODA currently uses the Merit Network Telescope [41], prior to January 2022, IODA used the UCSD Network Telescope [42]. IODA generates a Telescope alert when the current number of valid unique source IP addresses drops below 25% of the median of the preceding 7 day time-window. This threshold is significantly lower than the other signals, due to the higher variance of the Telescope signal.

3.1.2 Manually curated list of IODA outages. Since January 2018, we have used data from IODA's dashboard and API to identify and record the details of large-scale Internet outages. Typically, we identified potential countries, regions or ASes that might be experiencing outages by reviewing a list of the recently generated alerts on IODA's Outage Dashboard page [31]. In other cases, our investigations into potential outages via IODA's signals were initiated by requests for corroboration from other organizations or reading reports of Internet outages from other outage detection sources or news outlets (in most cases, IODA had also automated generated alerts that we had not yet noticed).

For an outage to be added to our list, the event must meet one of the two following requirements: *(i)* a prolonged (across multiple time bins) and significant drop is clearly visible in at least two of IODA's signals, with the drop in both signals overlapping temporally; or *(ii)* a prolonged and significant drop is visible in only one of IODA's signals but the outage has been corroborated by an external source, such as Kentik's Internet Outage Tracker [32] or Cloudflare Radar's Outage Center [15].

Once an outage is identified, the first two fields we populate are the start and end times (if the outage is ongoing we complete it once the outage has ended). In cases where two IODA signals demonstrate a disruption, the time of the first signal to drop is used as the start time of the disruption and the time of the last signal to recover is used as the end time. In the instances where an outage is visible in IODA and corroborated via external sources, we then ensure that the start and end times match a drop in at least one of IODA's signals. If there is a match, we use populate the time fields based on the signals in IODA. The date, hour, and minute are recorded for both start and end times.

The next field we record is the country/region/AS in which the outage took place. We then record the scope of the outage, which corresponds to the highest level of visibility of the outage. For example, if the outage is visible at the country level, we record "Country" as the scope. In instances where multiple countries experience an outage simultaneously (e.g., due to a cable cut), we record each of the affected countries individually as separate entries in our outage dataset. If we are unable to observe the outage at the country level, but do observe it for one or more sub-national regions, we record "Region" as the scope and record all affected regions. If the outage is only visible at the autonomous system level, we record "AS" as the scope. Our record for each outage also includes the list of signals

[2] Beginning September 2022, IODA integrated a fourth signal, the Google Transparency Report [25], into its country-level views. However, the data in our analysis predates this. As a result, we do not include a detailed description of this signal.

Start time	End time	Country	IODA BGP Auto Alert	IODA AP Auto Alert	IODA Telescope Auto Alert	IODA BGP visible by human	IODA AP visible by human	IODA Telescope visible by human	Scope	IODA URL	Cause	Confirmation Status	More Info
Thursday, Jun 30, 2022 5:30:00 AM	Thursday, June 30, 2022 10:40:00 PM	Sudan	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	Country	http://...	Government-ordered	Confirmed	Protests occurred; http://...

Table 1: An example of a record in our outage dataset corresponding to the event shown in Figure 1.

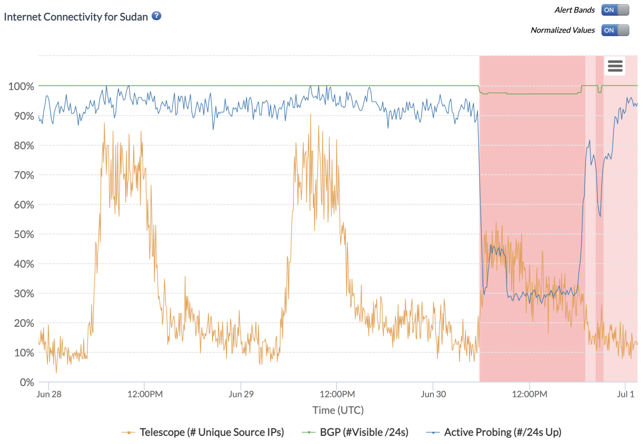


Figure 1: An example of IODA's view of an outage in Sudan during late June of 2022.

for which IODA generated automated alerts and the signals for which we manually observed a significant drop.

Once the scope and affected regions are recorded, we record the URL of the IODA page in which the outage is visible. If we are able to confirm the outage with an external source, we record this information as well, along with links to any discovered sources corroborating the event. In instances for which we are able to find additional reporting that provides further context relevant to the outage, such as related events or underlying causes (e.g., reports of protests or cable damage caused by damage during a natural disaster), we record links to these sources as well.

In certain situations, it is possible that dips in a signal are due to measurement artifacts or IODA infrastructure issues (e.g., an IODA measurement server failing, loss of BGP data due to a faulty collector, etc.). To account for such potential issues, prior to recording an observed drop as an outage, we first check other countries, regions, or ASes in different parts of the world as a control group. If we find that a similar drop is visible across disparate unrelated regions across the world, we do not record the outage in our dataset, as it is likely the result of an issue in IODA's infrastructure or its data sources and not the result of an actual outage.

Figure 1 depicts an example of IODA's view of an outage that occurred in Sudan in late June of 2022. Table 1 shows the corresponding entry in our outage dataset. In this example, the BGP and Active Probing signals drop at the same time, 5:30 AM UTC. Active Probing is the last signal to recover, so we use the recovery in Active Probing at 22:50 PM UTC as the end time. IODA generated an alert for both the BGP and Active Probing signals and we were able to manually observe the drop in the BGP, Active Probing, and Telescope signals. These details are recorded as true or false values in the appropriate fields. This particular outage was reported by the media and advocacy organizations to be a government-ordered shutdown during an active protest. We include these contextual details and links to them (truncated in the example) in the *Cause*, *Confirmation Status*, and *More Info* fields.

In addition to our ongoing, daily process of recording outages, we contracted a data service provider, called DataWorks[3], to review the historic data recorded in our manually curated list of IODA outages. The DataWorks team was hired and trained to review and add missing data fields including start and end times as well as which of IODA's signals demonstrated visible drops during an outage. A sample of their work was reviewed by our team to assure quality.

Due to issues in IODA's data collection and inconsistent investigation of outages, our curated list is not as comprehensive from August 1, 2021 to November 2021. Furthermore, the IODA website and measurement infrastructure service were offline for an extended period of time, from November 2021 to early February 2022, while it was migrated from one institution to another. Though we resumed consistent recording of outages in February 2022, our analysis in this paper incorporates other datasets that are published annually and were not yet available for 2022 at the time of our analysis. As such, we limit our period of study to outages occurring between January 1, 2018 and August 1, 2021. During this time period, we recorded a total of 896 country-level outage events across 155 countries.

3.2 Access Now #KeepItOn STOP

Access Now compiles the #KeepItOn STOP dataset (KIO) [2], a comprehensive dataset of network shutdowns and censorship events. This dataset contains events that involve an "intentional disruption of Internet or electronic communications, rendering them inaccessible or effectively unusable, for a specific population or within a location, often to exert control over the flow of information". The Access Now and the KIO dataset are important resources for insights on network shutdowns and censorship events for organizations such as the Freedom Online Coalition.

Access Now collects information from various sources including the civil society, governments, and corporate sources. The civil society sources are mainly composed of trusted local and international news media outlets, the Internet measurement experts, and local partners. Governments and corporations, such as telecommunications and social media companies, are other reliable sources of shutdowns and censorship events. In some circumstances, these entities publicly acknowledge their responsibility in conducting Internet shutdowns. Access Now also engages with volunteers inside governments and corporations that, under an agreement of anonymity, reveal information on shutdowns and censorship. To conceal the identity of these informants, Access Now does not include any label revealing the source of information.

[3] DataWorks is a data service provider that hires and upskills individuals from communities that have been historically excluded from computing.

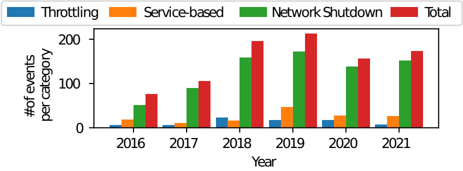


Figure 2: Number of KIO events per category and the total number of events per year.

Access Now relies on multiple sources to validate intention behind a shutdown. This validation process involves sources on the ground who may have experienced the shutdown. To ratify this information, Access Now further examines news media outlets, potential corporation sources, and members of the Internet measurement community. After gathering all this information, the event is added to the list if intentionality was detected with high confidence.

The KIO dataset also reports the type of restrictions applied during each event. There are three categories: *(i)* throttling, *(ii)* service-based bans and *(iii)* full-network shutdowns. Given that these categories are not mutually exclusive [48], events report a list of restrictions applied (e.g., throttling and service-based bans). The KIO dataset also details the regions where these restrictions took place. Though the KIO dataset includes many additional fields, the types of access networks affected (broadband networks, mobile networks, or both) and the geographic scope are the most relevant for the analysis in this work.

It is also important to note that a single entry in the KIO dataset can include multiple restrictions that occurred separately over the course of an event. As an example, a full network shutdown followed by an application ban will be listed as a single record, specifying that both a full network shutdown and service-based bans occurred during the event. It will not include, however, the specific time ranges during which each type of restriction was observed.

Similarly, a series of full-network shutdowns that have been mapped to a single overarching event will be listed as a single entry in the dataset. This includes exam-related shutdowns, such as those in in Iraq and Syria, as well as the shutdowns following the coup in Myanmar in 2021. In such cases, Internet shutdowns occurred during specific parts of the day across multiple weeks.

In this work, we use annual snapshots of the KIO dataset from 2016 through 2021. Access Now modified field names, value ranges, and the structure of the dataset several times during this period. We manually curated and homogenized the annual snapshots of the KIO dataset to facilitate the analysis in this work.

Figure 2 provides a summary of the types of techniques observed during events in the KIO dataset. As a reminder, the throttling, service-based, and full network shutdown categories are not mutually exclusive (multiple techniques might be used during a single event) and do not sum to the "Total". The plot shows that across all years of the KIO data, a large majority of the events involved full-network shutdowns. We also see that the number of full-network shutdowns grew significantly from 2016 to 2019. While censors are increasingly employing more sophisticated app-specific or throttling techniques, Figure 2 suggests that there are no signs of a significant decline in the occurrence of full-network shutdowns.

3.3 Additional datasets

The additional data sources in this work include countries' macroeconomic indicators, indices describing political regimes and freedom, and multiple Internet-related datasets. In our analysis, we leverage these datasets to conduct an exhaustive characterization of the circumstances under which Internet shutdowns take place.

Macroeconomic indicators. We use macroeconomic data gathered from the World Bank Data Bank (accessed 2023-01-23)[4], to investigate the influence of economic factors in political uses of Internet shutdowns. Our main measure of interest is gross domestic product (GDP) per capita and the prevalence of broadband Internet access. Our intuition follows the economics and information technology literature [13, 14, 21, 22] where trade, foreign direct investments (FDI) and economic growth are positively associated with increased Internet access.

Sociopolitical datasets. We use democracy scores from the Varieties of Democracy (V-Dem v11.1) database to identify democracy and autocracy levels of countries that engage in and experience Internet shutdowns [17]. While the Arab Spring serves as one of the earliest demonstrations of Internet shutdowns used by the state against the populace to minimize opposition mobilization [19], we have observed an increase over time in Internet shutdowns coinciding with political events; ranging from the coups and attempted coups in Ethiopia and Myanmar in 2019 and 2021 [48, 58], to the shutdowns in Gabon following the 2016 election [37]. We delve into this phenomenon using data on coups from the Global Instances of Coups dataset (accessed 2023-01-23) [50].[5] We manually collected data on elections between the years 2018-2021 from the International Foundation for Electoral Systems-Election Guide (accessed 2023-01-23)[6] as well as data identifying protest events using the Mass Mobilization in Autocracies Database (accessed 2023-01-23) [56]. Based on our experience observing Internet shutdowns, we expect there to be a significant relationship between Internet shutdowns and political events such as elections, protests and coups.

Computer network datasets. We also leverage multiple Internet-related datasets to investigate the relationship between the prevalence of state-ownership of the network and Internet shutdowns. We use two metrics to quantify the prevalence of states in domestic access markets, (1) the address space and (2) eyeballs. To calculate the domestic address space, we download CAIDA's daily prefix-to-AS mappings [10] and combine them with the MaxMind geolocation database [1]. This provides an estimate of the number of IP addresses per-AS in each country. To mitigate misrepresentation given the widespread use of NAT, we complement our analysis with population estimates for each AS. For this, we use APNIC's eyeball dataset (accessed 2022-03-21) [29], which uses an advertisement-based methodology to estimate user population. As a last step, we download a list of state-owned AS [11] to compute the prevalence of state-ownership in each country.

4 MERGED DATASETS

In this section, we describe our process for combining information across the sources described in Sec. 3 and provide a brief summary

[4] Data downloaded from https://databank.worldbank.org/.
[5] Data downloaded from https://arresteddictatorship.com/coups/.
[6] Data downloaded from https://www.electionguide.org/.

of the cleaned and merged dataset that we use throughout our analysis.

In order to combine multiple datasets, as a result of differences in naming conventions, we first needed to standardize the country names used across the datasets. These discrepancies were generally due to issues such as: using a different language (e.g., "Ivory Coast" vs "Cote d'Ivoire"), using names that had since changed (e.g., "Swaziland" vs "Eswatini"), containing minor differences in spelling (e.g., "Timor Leste" vs "Timor-Leste"), or using longer forms of a country's name (e.g., "Venezuela" vs "Venezuela, Bolivarian Republic of"). After standardizing the country names for the data, we used ISO 3166 2-letter alpha codes to identify each country across all datasets.

Our next goal was to identify events in the KIO and IODA datasets that corresponded to the same event. To do so, we matched all entries that were recorded in the same country during overlapping time periods. In the IODA data, each entry includes a start and end time in UTC. Events in the KIO dataset are recorded with a start and end date (localized to the country). Since the KIO entries do not include a time of day, when searching for a possible match, we assume 00:00:00 as the earliest possible start time and 23:59:59 as the latest possible end time, both in local time, for the dates specified.

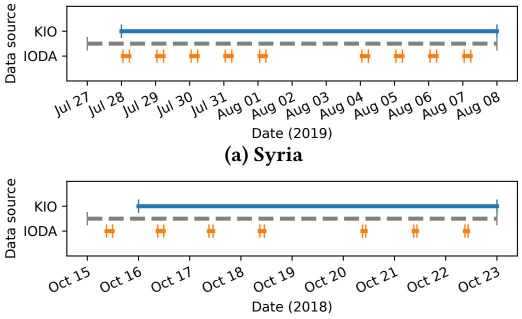
After converting the IODA start and stop times to local time (we use the timezone of country's capital city if the country spans multiple zones), we then matched an IODA event to a KIO event if its start time fell between the start and end of the KIO event.

In our initial inspection of the resulting matched dataset, we found instances of IODA outages that were not matched to a KIO event but appeared to be part of a series of shutdowns, the latter of which included IODA entries that matched a KIO event. This was due to the fact initial entry's start time in IODA was prior the start date listed in the KIO dataset, even after adjusting to local time.

Upon further investigation, we discovered that these discrepancies appeared to be the result of incorrect event details in the KIO dataset. This included mistakenly using the date on which the event was reported (i.e., the date of publication) rather than the date on which the event actually started. In such cases, we found the date listed in the KIO dataset matched the publication date of an article linked to in the event's description field, but not the start date reported in the article. Other causes included shutdowns starting close to midnight local time but not reported on until after midnight as well as articles using a timezone other than that of the country affected (e.g., local to the news organization's headquarters or their target audience).

To account for such issues, we expanded the time window for matching events to include the 24 hours preceding the KIO local start date. For each new match that we discovered by increasing the matching, we manually verified via news articles or communication with Access Now that IODA was accurately capturing the start time of the first relevant shutdown.

Figure 3 shows two examples of a KIO entry being matched to multiple IODA outages, one in Syria and one in Iraq. In each case, the single entry in the KIO dataset specifies the date range for the general event, while the IODA dataset provides details on the specific hours during which Internet access was shut down. We also note that the case in Fig 3b illustrates an example of the aforementioned issue where the first recorded IODA event occurred prior to the start date listed in the KIO dataset.



(a) Syria

(b) Iraq

Figure 3: Timeline of repeated full network shutdowns related to exams from 2019-07-28 to 2019-08-07 in Syria (top) and from 2018-10-15 to 2018-10-22 in Iraq (bottom). For each graph, the upper and lower lines represent the start and end times of individual entries in the KIO and IODA datasets, respectively. The gray dashed line in the middle represents the time ranges used for matching KIO and IODA events.

The primary advantages of the IODA dataset are that it provides a precise measurement of when the shutdown occurred as well as additional data at a technical level. On the other hand, the KIO dataset provides additional details on the intent and broader context of the shutdown we observed in IODA, simplifying the process of determining which shutdowns were related to the same phenomenon.

Shutdown and Outage Dataset. For our merged dataset of Internet shutdowns, we labeled the following events as *shutdowns*: (1) all KIO events that were identified as involving a full-network shutdown and (2) all IODA events that were either matched to a KIO event or were recorded by us as having a cause of either government-ordered or exam-related. All remaining IODA events (i.e., those that were not matched to a KIO event and listed a cause other than government-ordered or exam-related) were labeled as *spontaneous outages*. We again restrict our set of events to those that occurred between January 1, 2018 and August 1, 2021, the time period for which data from KIO and IODA overlap.

For our study in this paper, we filter our merged dataset such that it only includes shutdowns and spontaneous outages that were observed at the country level. Although our outage dataset does includes disruptions that occurred at the regional and AS levels, we choose to focus on country-level events for multiple reasons.

The biggest barrier to incorporating regional shutdowns into our analysis is the result of two compounding issues. First, is that shutdowns that occur at the subnational level are very highly concentrated; according to the KIO dataset, we find that 85% of subnational full-network shutdowns occur in India. For 72% of these events, only mobile networks were affected. The second related issue is that although IODA is able to monitor publicly routable IPv4 networks, its ability to monitor the connectivity of networks that heavily utilize Network Address Translation (NAT), such as mobile networks, is limited. As a result, IODA's set of shutdowns at the subnational level would lack details on a significant fraction of such events.

Furthermore, the macroeconomic and sociopolitical index datasets used in our analysis are only available at the country level for each year, limiting our ability to study the relationship between economic and sociopolitical characteristics specific to the regions affected by shutdowns.

Table 2 shows a break down of the number of country-level shutdowns and spontaneous outages in our resulting dataset. It also lists the top five countries in terms of highest number of events for each category. Our final dataset contains a total of 219 national-scale Internet shutdowns in 35 countries and 714 spontaneous outages in 150 countries. Using DataReportal's estimated number of Internet users per country [20], the 35 countries experiencing Internet shutdowns together represent an estimate of more than 1 billion Internet users.

5 DATA ANALYSIS

To date, political science research has mostly focused on the institutional and event correlates of Internet shutdowns. In particular, these studies have noted that countries that are more authoritarian, have fewer media freedoms and more corruption, are more likely to have shutdowns [28]. More recent work has shown that state ownership of the ISP space and elections in authoritarian regimes also predict shutdowns [23].

Analyzing shutdowns exclusively does not allow us to distinguish between shutdowns and spontaneous Internet outages. Authoritarianism also correlates with low investment in infrastructure and lower GDP, both which we would expect would be correlated with spontaneous outages [3, 7, 36]. When an outage occurs, to what extent do economic indicators and the political institutions of the country predict whether the outage was political in nature? Which types of predictors tend to be most reliable in distinguishing between shutdowns and spontaneous outages?

In this section, we reanalyze the institutional correlates of national shutdowns and outages, but use the merged KIO and IODA datasets to compare political shutdowns to non-political spontaneous outages. We find that indicators of non-democracies, including regime type, military control of the regime, and media corruption and bias are not only associated with shutdowns, but also associated with spontaneous outages as well. Interestingly, political events such as elections, protests, and coups are predictive of shutdowns, but not of spontaneous outages. In addition, state control of the address space tends to predict shutdowns, but not spontaneous outages. We find a variety of other technical indicators, including the timing, length, and recurrence of the outage may also provide reliable indicators of shutdowns.

5.1 Political Institutions and Outages

We first investigate the extent to which political institutions are indicators of the political nature of an Internet outage. Scholars have long shown a robust link between Internet censorship and authoritarianism [28], including non-representative political institutions, military control of the government and media control. Yet, authoritarianism is also associated with lower GDP, worse public goods provision and failing infrastructure [7, 36], which may also be predictive of spontaneous outages.

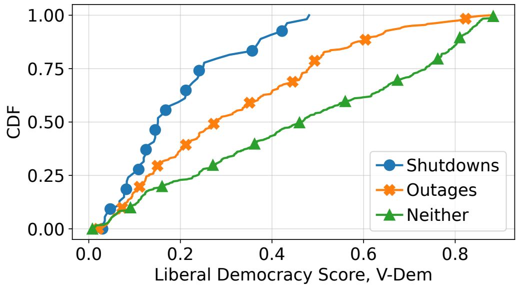


Figure 4: Autocracies are more likely to have both shutdowns and spontaneous outages, though the difference is more stark with shutdowns.

We use the V-Dem dataset described in Section 3.3 to evaluate how well political institutions might be a useful indicator for distinguishing shutdowns and spontaneous outages. To do this, we first identify all the country-year combinations during which a country experienced at least one national-scale shutdown for a given year. In a similar manner, we then identify the country-year combinations with at least one national-scale spontaneous outage. We then assign each country-year to a group: those with shutdowns ("Shutdowns"), those with spontaneous outages ("Outages"), or those with neither ("Neither"). Table 3 shows the total number of country-year combinations assigned to each category.

Note that because our analysis is at the country-year level, if a country has a shutdown in one year and neither a shutdown nor a spontaneous outage in another year, the same country will appear in both the "Shutdowns" and "Neither" categories (except for different years). As an example, we did not observe any spontaneous outages or shutdowns in Myanmar during 2018, but did observe shutdowns in all other years of our study. Thus, "Myanmar-2018" is categorized under "Neither" while all other years are included under "Shutdowns". We then merged this data with indicators of political institutions for each country-year.

Figure 4 shows the distribution of liberal democracy scores provided by V-Dem for each category. For this variable, a lower scores represents a more autocratic government. We observe that while there is overlap between the distributions, there is a clear difference in the overall distribution of scores across each group. Overall, country-year combinations with shutdowns have the lowest scores (median of 0.151, maximum of 0.481). Interestingly, we find that country-year combinations with spontaneous outages are also likely to be more authoritarian (median of 0.279) in comparison to countries that experience neither (median of 0.465).

These patterns also appear in other political variables from the V-Dem dataset, such as in V-Dem's measure of the military's capability to remove the existing regime (higher values suggest the military is more capable of removing the existing regime). Figure 5 shows the distribution of the political power of the military across each of the three categories, where over half of the country-year combinations that experienced neither class of event have scores of 0, with median scores increasing to 0.25 for those that experienced spontaneous outages and 0.33 for those that experienced shutdowns.

Figure 6 shows that both shutdowns and spontaneous outages have some degree of association with V-Dem's measures of media bias and freedom of expression among men. Both of these variables

	Country-level shutdown events				Country-level spontaneous outage events
	KIO		IODA		IODA
	Total #	# matched to IODA	Total #	# matched to KIO	Total #
Event count	82	45	182	152	714
Top 5 countries by # of events	Iraq (14) Myanmar (7) Algeria (6) Syria (5) Iran (4)	Iraq (7) Myanmar (6) Syria (5) Iran (3) Ethiopia (3) Algeria (3)	Myanmar (53) Syria (52) Iraq (38) Eswatini (5) Ethiopia (5)	Myanmar (53) Syria (43) Iraq (22) Eswatini (5) Ethiopia (5)	Togo (40) Venezuela (36) Niger (23) Eswatini (20) Cameroon (19)

Table 2: Summary of the number of country-level shutdown and outages events per category in the merged KIO-IODA dataset. The bottom row lists the top five countries (more in cases of a tie) in terms of the number of events in per category. Our merged dataset of shutdowns contains: (1) all KIO events identified as involving a full-network shutdown and (2) all IODA events that were either matched to a KIO event or were recorded by us as having a cause of either government-ordered or exam-related.

Country-years w/ Shutdowns	Country-years w/ Outages	Country-years w/ Neither
55	310	514

Table 3: Summary of the number of country-years combinations in each category.

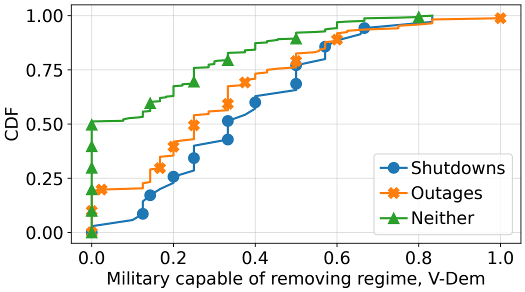


Figure 5: Both shutdowns and spontaneous outages are more likely in country-years where military is politically powerful.

are based on scores collected via surveys. V-Dem transforms their values such that 0 approximately represents the mean of all country-years (though this is not necessarily normally distributed). Lower values represent a higher degree of authoritarianism (higher median bias and less freedom of discussion). Across both variables, countries with shutdowns tend to be more associated with a slightly higher degree of authoritarianism. Countries with spontaneous outages also tend to have more media bias and less freedom of discussion for men than countries that experience neither types of events.

Why are countries that experience spontaneous outages more likely to have authoritarian institutions and media bias than those who experience neither spontaneous outages nor shutdowns? One possible explanation is that countries that have more authoritarian institutions are also more likely to have lower GDP and underinvest in public goods, meaning less diversified or more centralized infrastructure and thus more prone to failure. Indeed, we also find this correlation with shutdowns and spontaneous outages, as presented in Figure 7. Countries that shutdown the Internet are most likely to have low GDP per capita and limited broadband access. Countries that experience spontaneous outages are also more likely to have low GDP per capita and limited broadband access, though to a lesser extent.

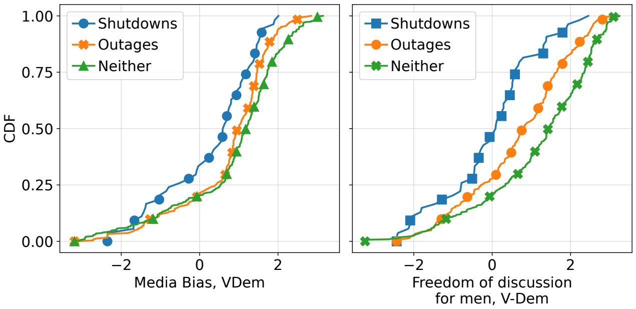


Figure 6: Both shutdowns and spontaneous outages are more likely in country-years with media bias and less freedom of discussion.

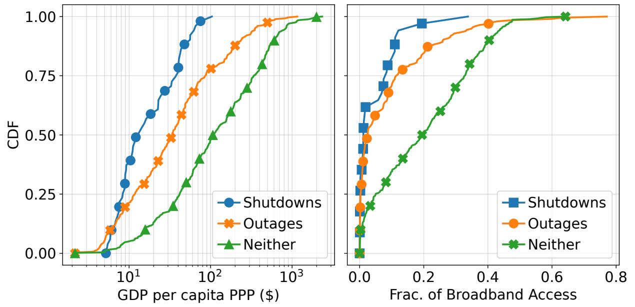


Figure 7: Both shutdowns and spontaneous outages are more likely in country-years with low GDP and less access to broadband.

5.1.1 State ownership of the domestic address space. Next, we incorporate the state ownership of Internet operators as a factor to predict shutdowns. State-ownership of network operators[7] provides governments with both control of the company and the direct means to execute Internet disruptions.

While governments may be involved a wide range of Internet services, we focus on participation in the domestic access market. As last-mile providers, governments are in full control to disconnect users from the network. The effectiveness of this disconnection mechanism relies on the assumption that users are not going to have backup connections to get back online.

[7] We consider state-owned operators to be those that are controlled by the government through the ownership of more than 50% of the shares.

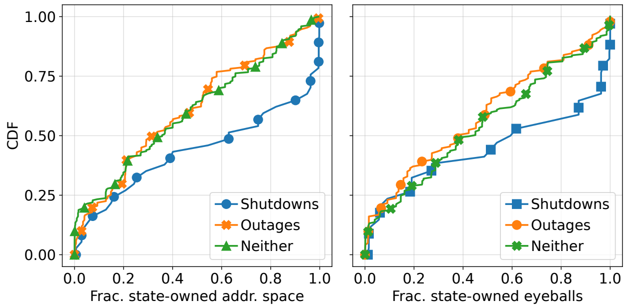


Figure 8: Shutdowns are more prevalent in countries with higher presence of the state in the address space and eyeballs.

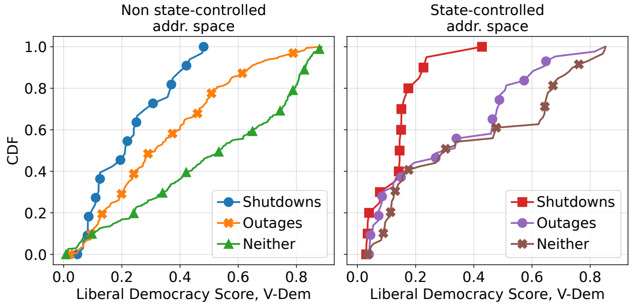


Figure 9: Both shutdowns and spontaneous outages are more prevalent in countries with low liberal democracy scores that originate the majority of the domestic address space.

We investigate whether the prevalence of the government in the access market correlates with shutdowns. We use the fraction of domestic address space and the fraction of eyeballs connected via state-owned operators as indicators of the government's participation in access provisioning. Figure 8 shows the distribution of these two variables across each group in years known to include state-owned providers. While we observe no discernible difference between the spontaneous outage and neither categories, the curve for shutdowns shows a different behavior. Concretely, in countries with state-owned providers, shutdowns are more prevalent where there is larger control of the address space or eyeballs.

We further investigate whether V-Dem's liberal democracy score is a predictor of shutdowns in countries where the state hold the majority of the address space. For this analysis, we convert the state-ownership of the address space into a categorical variable. For this, we define a state-controlled address space as a country where state-owned providers originate more than 50% of the domestic address space.

Figure 9 shows the distribution of the liberal democracy scores across the three categories. The figure splits the analysis across two plots, showing countries without state-control of the address space on the left and countries with state-controlled address space on the right.

Comparing both panels, we observe that the curve of shutdowns is skewed to the left in countries with state-controlled address space. Indeed, the mean value of the Liberal Democracy score is 0.13 and 0.22 in countries controlling and non-controlling the address space, respectively. This indicates that in countries with state control of the address space, authoritarianism is a much better indicator of a shutdown than in countries that do not control the address space.[8]

Event	Pr(Shutdown)	Pr(Outage)
Election	0.016	0.003
No Election	0.001	0.002
Coup	0.286	0.000
No Coup	0.001	0.002
Protest	0.009	0.004
No Protest	0.001	0.002

Table 4: The probability of shutdowns and spontaneous outages on days where that country has an election, coup, and protest in comparison to days where the country does not have an election, coup or protest.

5.2 Mobilization events predict shutdowns

In this section, we investigate whether political events predict shutdowns and whether or not these same events also predict spontaneous outages. It is well known that censorship has been used as a strategy for controlling election outcomes [23, 38], preventing protests [16, 33, 34, 54], hiding human rights abuses [24], and manipulating coup outcomes [8, 48]. While both spontaneous outages and shutdowns are associated with authoritarian institutions, if spontaneous outages are also associated with major political events such as elections, coups, and protests, we might wonder whether such events are indeed political, going unidentified as shutdowns.

To investigate this, we use the data described in Section 3.3 on elections, coups, and protests from countries around the world during the same time period as our merged dataset. We then estimate the overall probability of a shutdown occurring on the same day in the same country as an election, coup, and protest, in comparison to days where that country had no elections, no coups and no protests.[9]

Our results are reported in Table 4. We find significantly higher likelihoods of Internet shutdowns on days where the country has an election, coup, or protest. An election increases the probability of a shutdown by an order of 16, a coup by almost an order of 300, and a protest by an order of 9.[10] In contrast, these events are not associated with an increased probability of a spontaneous outage.[11]

5.3 Technical indicators of shutdowns

We now analyze the temporal and technical characteristics of Internet shutdowns in comparison to spontaneous outages. Our analysis covers temporal characteristics, such as duration, recurrence rate, and start times and finds that shutdowns differ significantly compared to spontaneous outages. These differences suggest that the

[8] Though not included here, we observed similar trends when using the eyeball dataset to quantify state-ownership of the address space.

[9] For protests, we only have data through 2019, so we subset our analysis to 2018 and 2019.

[10] Note that there are only seven coups in the dataset.

[11] Our analysis is robust to several different approaches. Considering within-country trends in shutdowns and spontaneous outages (including country and day fixed effects), these results still hold. Aggregating to the week level instead of the day level also produces the same results.

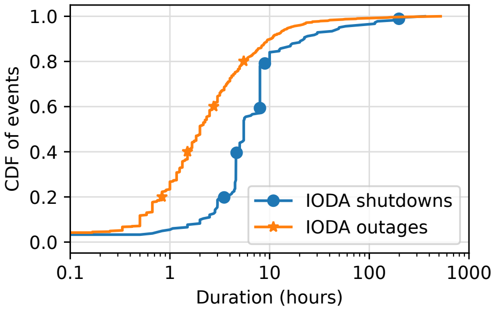


Figure 10: CDF of the durations of shutdowns and spontaneous outages observed via IODA.

events categorized as shutdowns are indeed the result of intentional human interventions (i.e., government-ordered). Such evidence could be beneficial to Internet freedom advocacy and human rights organizations. Even in cases where a government later admits to executing an Internet shutdown, understanding the unique characteristics of Internet shutdowns could help rapid response efforts differentiate between shutdowns and spontaneous outages.

For this analysis, we focus specifically on the events listed in our IODA dataset. We split these events into two groups based on whether or not we had previously identified the event to be connected to an Internet shutdown.

The first group is composed of all IODA events identified as shutdowns. Recall that an IODA event was tagged as a shutdown if the event matched to a KIO shutdown (i.e., occurred simultaneously in the same country) or had been identified by us as being caused by a shutdown (i.e., our entry was annotated as either government-ordered or exam-related based on news and other reporting at the time). The analysis in this section refers to these events as the "IODA shutdowns" set and includes 182 events in 24 countries (corresponding to the center columns of Table 2). Of these, 133 had been identified as a shutdown event by both matching to a KIO event as well as being identified by us as a shutdown. The remaining events were identified as shutdowns as a result of only matching to a KIO event (19) or only being identified by us as a shutdown (30).

The second group is composed of the remaining events, (i.e., those that were not identified as shutdowns). We consider these to be spontaneous outages and refer to them as the "IODA outages" set in our analysis below. This set contains 729 outages across 151 countries.

One of the main advantages of the IODA dataset is that it provides an objective, technical measure of Internet service availability. In contrast to the KIO dataset, this allow for a fine-grained timeline of the status of Internet connectivity. Since our analysis in this section focuses on the fine-grained temporal and technical aspects of Internet shutdowns and spontaneous outages, we do not include KIO events in this analysis (though we do retain IODA events tagged as shutdowns using KIO data).

Event duration. We first look at how shutdowns and spontaneous outages compare in terms of duration. Figure 10 shows the CDF of event durations for each category. We find that spontaneous outages tend to have shorter durations, with a median duration of 2 hours for spontaneous outages and 5.5 hours for shutdowns. We also find that shutdowns are significantly more likely to have a duration that is a multiple of 30 minutes, with over 55% of shutdowns lasting a multiple of 30 minutes, compared to 15% of spontaneous outages. We also find that a particularly high fraction of shutdowns last precisely 4.5, 5.5, 8, or 10 hours (45%), compared to less than 1% of spontaneous outages with those same durations.

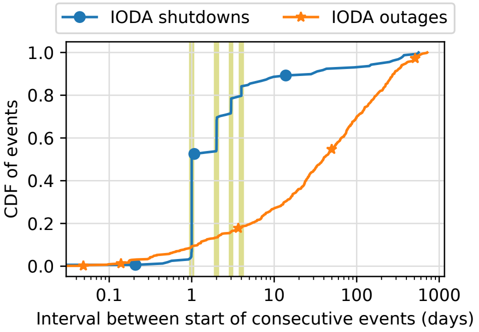


Figure 11: Time interval between the start times of consecutive shutdown and spontaneous outage events within the same country.

Recurrence interval. We next look at the recurrence of outage events for each category. Of the 24 countries that had an Internet shutdown between January 1, 2018 and August 1, 2021, 50% (12) experience a second shutdown in the same time period. Surprisingly, we find that countries that experienced spontaneous outages, were actually more likely to see a subsequent spontaneous outage, with at least a second outage occurring in 72.2% of the 151 countries with spontaneous outages.

Though countries with shutdowns were less likely to have recurring shutdowns compared to spontaneous outages over the time period of our study, recurring shutdowns happen at significantly shorter intervals. Figure 11 shows the interval between successive start times for shutdowns and spontaneous outages. The median interval between shutdown events within the same country was 1 day for shutdowns compared to 39 days for spontaneous outages. Additionally, the distribution for shutdowns is largely concentrated on a small number of specific values (similar to Figure 10). The vertical bars in Figure 11 highlight intervals of exactly 1, 2, 3, or 4 days. We find that 67.7% of all shutdowns fall precisely in these time intervals, compared to just 0.17% of all spontaneous outages.

Start times. The last of the temporal characteristics we analyze involve the start times of events. For this, we examine how event start times are distributed across the minutes of the hour, the hours of the day, and the days of the week.

Figure 12 shows the distribution of the minute of the hour we recorded as the start time for shutdown and spontaneous outage events (we recorded event start times in UTC). Note that due to the difference in time-granularity of the IODA signals (having both 5- and 10-minute intervals), events are more likely to be recorded as starting on minutes that are multiples of 10, as those time bins have data for all three signals rather than just two. Overall, we find that the percentage of shutdowns that start on either the hour or half hour is much higher compared to spontaneous outages, at 87.4% and 39.6%, respectively.

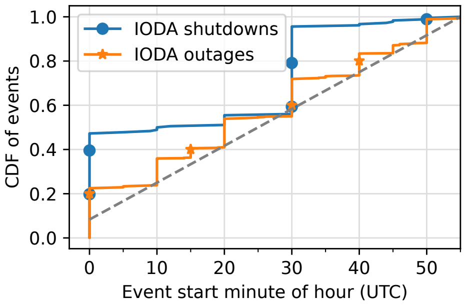


Figure 12: Minute of the hour (UTC time) of start times for shutdowns and outages recorded via IODA. For comparison, the diagonal dashed gray line represents a uniform distribution across each 5-minute bucket.

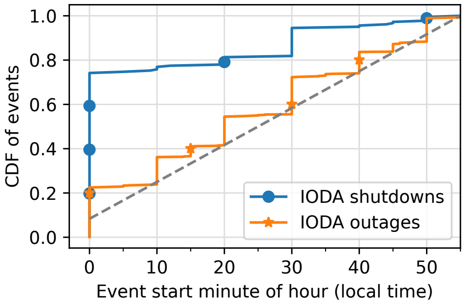


Figure 13: CDF of shutdown and spontaneous outage start times by minute of the hour (local time). The diagonal dashed gray line represents a uniform distribution across each 5-minute bucket.

Due to the fact that a number of timezones include a half hour offset in their minute component, we also calculate the starting minute for each event in local time, shown in Figure 13. In cases where a country spans multiple timezones, we use the timezone of that country's capital city as an estimate of the appropriate timezone. Figure 13 shows that the number of shutdowns that begin on the hour after converting to local time increases from 47.3% to 74.2%. Interestingly, we also observed that in contrast to shutdowns, the conversion to local time appears to have no noticeable impact on the distribution of spontaneous outages across the hour. In both cases, the distribution of spontaneous outages is similar to a uniform distribution across the 12 possible time buckets (represented by the diagonal dashed gray line).

Next, we look at at the distribution of event start times (local time) across the hours of the day. We find that a disproportionate number of shutdown events starting on specific hours, with 72.1% of shutdowns starting between 00:00 and 06:00 (inclusive, local time). Shutdown events across these hours were largely weighted by practices in specific countries, including nightly shutdowns starting at 00:00 in Myanmar in 2021 and exam-related shutdowns in Syria across multiple years starting at 02:00 and 04:00.

Shutdowns also differ in comparison to spontaneous outages in terms of the day of the week on which they began. Figure 15 shows the distribution of shutdowns and spontaneous outages according to the weekday that the event started (after converting to local time). Overall, spontaneous outages tend to follow a fairly uniform distribution (represented by the horizontal dashed gray line) across the days of the week. However, shutdowns are noticeably less likely to occur on Fridays and, to a less extent, Saturdays. Using a two-tailed binomial test across both sets of events and days of the week, we found that the lower number of shutdowns on Fridays was a statistically significant deviation from the expected uniform distribution (p-value < 0.00065).

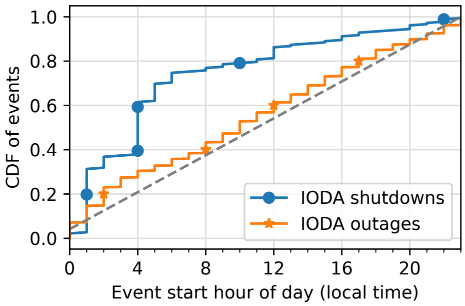


Figure 14: CDF of nationwide shutdown and spontaneous outage events according to hour of the day (local time). The diagonal gray dashed line represents a uniform distribution across the day.

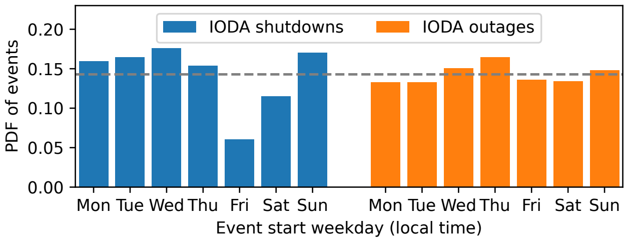


Figure 15: Day of the week during which IODA shutdowns and spontaneous outages started (in local time). The dashed gray line represents a uniform distribution ($y = 1/7$).

It is worth noting that a number of regions and cultures do not include Friday as part of the customary workweek, including regions in Syria, Iraq, Iran, Sudan, and Algeria, which together account for 57% of the shutdowns in this set.

We expect that classifying each day of the week as either a local customary workweek or weekend would further highlight the disproportionate concentration of events onto working days. Unfortunately, we were unable to find a reliable dataset containing workweek customs across all countries in our dataset. This is a particularly difficult categorization to make as it can even vary across regions within a single country and also be specific in each region to a sector (public or private). The readily available datasets that we did find tended to lack coverage, particularly in the global south, which represents a majority of all shutdowns. In some cases, we were able to find auxiliary sources with information on local customs in regions that lacked coverage, however, investigating these sources further we found multiple disagreements depending on the source.

Overall, analyzing the time-related characteristics of Internet shutdowns in comparison to spontaneous outages shows that shutdowns deviate remarkably compared to spontaneous outages, which

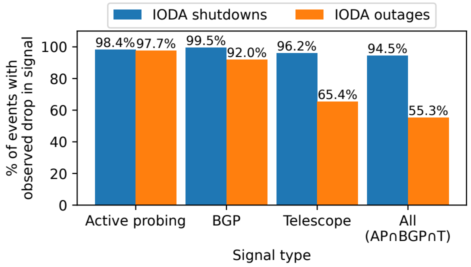


Figure 16: The percentage of shutdowns and outages for which we observed a significant drop for each signal.

tend to be closer to a uniform random distribution. This is in line with the planned and unplanned nature of each category of event.

IODA signals with observable drops. Our final comparison of shutdown and spontaneous outage events compares which of IODA's signals are likely to have had an observable drop. Figure 16 shows, for each time series signal, the percentage of events of each type for which we recorded an observable drop. The "All" category represents the percent of events in each category for which the drop was observed across all three metrics.

We find that, overall, shutdowns are much more likely to be observable across all three signals, with 94.5% of all shutdowns displaying a drop in all three signals. In contrast, spontaneous outages were much less likely to be observed across all three signals at 55.3%. This is largely due to the fact that IODA's Telescope signal was less likely to show a discernible drop for spontaneous outages. This suggests that shutdown events are more likely to result in a drop across all three of IODA's signals.

The results presented throughout this section show that Internet shutdowns have a number of unique characteristics compared to spontaneous outages across multiple economic, social, and technical indicators.

6 ETHICAL CONSIDERATIONS

Though our process of collecting data from the IODA dashboard and API did not involve human subjects, there are still ethical considerations in this work. First, considering that this research seeks to provide empirical data on the distinct characteristics of shutdowns versus outages, this research could provide censors with insights on how to evade shutdown measurements. However, our findings can be used to help better prepare advocacy organizations, litigation teams, and intergovernmental organizations, with insights into how to better identify when and where shutdowns occur and provide interventions of censorship evasion, negotiations, and litigation. As such, we believe that concerns related to publishing this information are outweighed by the benefit of increased public awareness and understanding of Internet shutdowns.

Our analysis incorporated AccessNow's publicly available KIO dataset, which does include shutdown and censorship events reported by human subjects. We believe that AccessNow has taken responsible steps to protect its sources' anonymity and minimize the risk of potential harm. In cases where events are reported by sources that are at personal risk or when sources who work inside government or corporations, AccessNow only reports the details of the shutdown and does not disclose details on their source.

7 CONCLUSION

This paper presents the first interdisciplinary, empirical analysis of longitudinal data on national-scale Internet shutdowns alongside non-political spontaneous Internet outages. Analyzing these two types of Internet disruptions together brings greater insight into the defining characteristics of Internet shutdowns, thus allowing us to better identify the signatures that distinguish Internet shutdowns from spontaneous outages.

Furthermore, understanding the differences between Internet shutdowns and spontaneous outages provides insight into the nature and origins of shutdowns. Surprisingly, we find that while institutional variables such as authoritarianism, media freedom, and economic development are associated with shutdowns, they are also associated with spontaneous outages. This may reflect a wider lack of investment in infrastructure and the economy among countries with high authoritarianism and low media freedom, conditions that precipitate spontaneous outages as well as more politically-motivated shutdowns.

Future Work. Because the existence and rationale of a shutdown are not typically disclosed, our findings could provide useful indicators for policymakers and those aiming to identify the source and nature of an outage and quickly put pressure on governmental institutions looking to cut off Internet access for political reasons. We plan to use our analysis to inform the design of a shutdown identification tool with heuristics of what to look for when assessing a disruption as a potential shutdown. With such a tool, the Internet freedom community could have greater confidence in identifying the source or nature of an outage. Examples of the questions that could be presented to investigators include: (1) *Did the disruption occur in a country that is an autocracy?* (2) *Did the disruption co-occur with an election, coup, or protest?* (3) *Did the disruption start on the hour in local time?* (4) *Did all three of IODA's signals simultaneously drop during the disruption?* Such a heuristic may not be perfect, but could help inform the allocation of investigatory resources in the immediate aftermath of a disruption.

We are also exploring the feasibility of using our findings to create a classifier for rapid and automated identification of Internet shutdowns. However, the economic, sociopolitical, and event data used in our dataset are limited in their frequency of updates, which are not available on an immediate basis. Though some indicators, such as GDP and V-Dem indicies, might typically be stable, this is not necessarily the case when a government's shutdown behavior changes drastically (e.g., shutdowns surrounding widespread protests, military coups, disputed elections, etc.). In future work, we plan to explore ways of addressing these limitations while ensuring that such a classifier does not lead to inaccurate reporting of Internet shutdowns.

ACKNOWLEDGEMENTS

We thank our shepherd, Suman Banerjee, and the anonymous reviewers for their thoughtful feedback. We also thank Venise Ivey, Jessica Hill, Justin Booker, and Dana Priest from Dataworks and Mehmet Islamoglu for their contributions in validating our IODA outage dataset.

REFERENCES

[1] 2023. Maxmind Geolocation Data. https://www.maxmind.com/en/geoip2-services-and-databases. (2023). Accessed: 2023-2-8.

[2] AccessNow. [n. d.]. #KeepItOn: Fighting internet shutdowns around the world. https://www.accessnow.org/keepiton/. ([n. d.]). Accessed: 2023-2-15.

[3] Daron Acemoglu, Suresh Naidu, Pascual Restrepo, and James A Robinson. 2019. Democracy does cause growth. *Journal of political economy* 127, 1 (2019), 47–100.

[4] Simurgh Aryan, Homa Aryan, and J. Alex Halderman. 2013. Internet Censorship in Iran: A First Look. In *USENIX Workshop on Free and Open Communications on the Internet*. USENIX Association, Washington, D.C. https://www.usenix.org/conference/foci13/workshop-program/presentation/aryan

[5] Yohannes Eneyew Ayalew. 2019. The Internet shutdown muzzle(s) freedom of expression in Ethiopia: competing narratives. *Information & Communications Technology Law* 28, 2 (2019), 208–224. https://doi.org/10.1080/13600834.2019.1619906

[6] Simone Basso, Maria Xynou, Arturo Filastò, and Amanda Meng. 2022. Iran blocks social media, app stores and encrypted DNS amid Mahsa Amini protests. https://ooni.org/post/2022-iran-blocks-social-media-mahsa-amini-protests/. (2022). Accessed: 2023-2-15.

[7] Matthew A Baum and David A Lake. 2003. The political economy of growth: democracy and human capital. *American journal of political science* 47, 2 (2003), 333–347.

[8] Raphael Boleslavsky, Mehdi Shadmehr, and Konstantin Sonin. 2021. Media Freedom in the Shadow of a Coup. *Journal of the European Economic Association* 19, 3 (2021), 1782–1815.

[9] CAIDA. [n. d.]. BGPView. https://github.com/CAIDA/bgpview. ([n. d.]). Accessed: 2023-2-8.

[10] CAIDA. 2021. RouteViews IPv4 Prefix to AS mappings. https://catalog.caida.org/details/dataset/routeviews_ipv4_prefix2as. (2021). Accessed: 2022-8-10.

[11] Esteban Carisimo, Alexander Gamero-Garrido, Alex C. Snoeren, and Alberto Dainotti. 2021. Identifying ASes of State-Owned Internet Operators. In *Proc. of IMC*.

[12] Censored Planet. 2023. Censored Planet: An Internet-wide, Longitudinal Censorship Observatory. https://censoredplanet.org/. (2023). Accessed: 2023-2-15.

[13] Changkyu Choi. 2003. Does the Internet stimulate inward foreign direct investment? *Journal of Policy Modeling* 25, 4 (2003), 319–326.

[14] Changkyu Choi. 2010. The effect of the Internet on service trade. *Economics Letters* 109, 2 (2010), 102–104.

[15] Cloudflare. 2023. Cloudflare Radar. https://radar.cloudflare.com/. (2023). Accessed: 2023-2-8.

[16] Michael Collyer and Joss Wright. 2021. A Bayesian Analysis of Collective Action and Internet Shutdowns in India. In *13th ACM Web Science Conference 2021*. 309–318.

[17] Michael Coppedge, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Nazifa Alizada, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Allen Hicken, Garry Hindle, Nina Ilchenko, Joshua Krusell, Anna L uhrmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Ei tan Tzelgov, Yi ting Wang, Tore Wig, Steven Wilson, and Daniel Ziblatt. 2021. V-Dem Country-Year/Country-Date Dataset v11. (2021). https://www.v-dem.net/en/data/data-version-11/

[18] Alberto Dainotti, Karyn Benson, Alistair King, kc claffy, Michael Kallitsis, Eduard Glatz, and Xenofontas Dimitropoulos. 2014. Estimating Internet Address Space Usage through Passive Measurements. *ACM SIGCOMM Computer Communication Review* 44, 1 (dec 2014), 42–49. https://doi.org/10.1145/2567561.2567568

[19] Alberto Dainotti, Claudio Squarcella, Emile Aben, Kimberly C Claffy, Marco Chiesa, Michele Russo, and Antonio Pescapé. 2011. Analysis of country-wide internet outages caused by censorship. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. 1–18.

[20] DataReportal. [n. d.]. Digital Around the World. https://datareportal.com/global-digital-overview. ([n. d.]). Accessed: 2023-2-15.

[21] Caroline Freund and Diana Weinhold. 2002. The Internet and international trade in services. *American Economic Review* 92, 2 (2002), 236–240.

[22] Caroline L Freund and Diana Weinhold. 2004. The effect of the Internet on international trade. *Journal of international economics* 62, 1 (2004), 171–189.

[23] Tina Freyburg and Lisa Garbe. 2018. Blocking the bottleneck: Internet shutdowns and ownership at election times in sub-Saharan Africa. *International Journal of Communication* 12 (2018), 3896–3916.

[24] Anita R Gohdes. 2015. Pulling the plug: Network disruptions and violence in civil conflict. *Journal of Peace Research* 52, 3 (2015), 352–367.

[25] Google. 2023. Google Transparency Report. https://transparencyreport.google.com/?hl=en. (2023). Accessed: 2023-2-8.

[26] Google. 2023. Jigsaw. https://jigsaw.google.com/. (2023). Accessed: 2023-2-13.

[27] Marianna Diaz Hernandez, Felicia Anthonio, Sage Cheng, and Alexia Skok. 2022. Internet shutdowns in 2021: the return of digital authoritarianism. https://www.accessnow.org/internet-shutdowns-2021/. (2022).

[28] Philip N Howard, Sheetal D Agarwal, and Muzammil M Hussain. 2011. When do states disconnect their digital networks? Regime responses to the political uses of social media. *The Communication Review* 14, 3 (2011), 216–232.

[29] Geoff Huston. 2014. How Big is that Network? https://labs.apnic.net/?p=526. (2014).

[30] Internet Intelligence Lab, Georgia Tech. [n. d.]. IODA (Internet Outage Detection and Analysis). https://ioda.inetintel.cc.gatech.edu/project. ([n. d.]). (Previously http://ioda.caida.org UC San Diego CAIDA.) Accessed: 2023-02-08.

[31] IODA. [n. d.]. Dashboard for Monitoring Internet Outages. https://ioda.inetintel.cc.gatech.edu/dashboard. ([n. d.]). Accessed: 2023-2-8.

[32] Kentik. [n. d.]. Internet Outage Tracker and Network Analysis Center. https://www.kentik.com/analysis/. ([n. d.]). Accessed: 2023-2-8.

[33] Gary King, Jennifer Pan, and Margaret E Roberts. 2013. How censorship in China allows government criticism but silences collective expression. *American political science Review* 107, 2 (2013), 326–343.

[34] Gary King, Jennifer Pan, and Margaret E Roberts. 2014. Reverse-engineering censorship in China: Randomized experimentation and participant observation. *Science* 345, 6199 (2014), 1251722.

[35] Georgia Tech's Internet Intelligence Lab. 2023. IODA Help page. https://ioda.inetintel.cc.gatech.edu/help. (2023). Accessed: 2023-2-8.

[36] David A Lake and Matthew A Baum. 2001. The invisible hand of democracy: political control and the provision of public services. *Comparative political studies* 34, 6 (2001), 587–621.

[37] Selena Larson. 2016. This African country is taking an unprecedented step in internet censorship. *CNN Business* (Sep 2016).

[38] Philipp M Lutscher, Nils B Weidmann, Margaret E Roberts, Mattijs Jonker, Alistair King, and Alberto Dainotti. 2020. At home and abroad: The use of denial-of-service attacks during elections in nondemocratic regimes. *Journal of Conflict Resolution* 64, 2-3 (2020), 373–401.

[39] Alessandro Mascellino. 2023. Over Four Billion People Affected By Internet Censorship in 2022. https://www.infosecurity-magazine.com/news/four-billion-people-internet. (2023).

[40] Mozilla Telemetry Data. 2023. Mozilla Telemetry Data. https://firefox-source-docs.mozilla.org/toolkit/components/telemetry/index.html. (2023). Accessed: 2023-2-15.

[41] Merit Network. 2023. The ORION NETWORK TELESCOPE. https://www.merit.edu/initiatives/orion-network-telescope/. (2023). Accessed: 2023-2-8.

[42] University of California San Diego. 2023. UCSD Network Telescope. https://www.caida.org/projects/network_telescope/. (2023). Accessed: 2023-2-8.

[43] University of Oregon. 2023. RouteViews. http://www.routeviews.org/routeviews/. (2023). Accessed: 2023-2-8.

[44] Office of the high commissioner for human rights. 2022. Internet shutdowns: UN report details 'dramatic' impact on people's lives and human rights. https://www.ohchr.org/en/press-releases/2022/06/internet-shutdowns-un-report-details-dramatic-impact-peoples-lives-and-human. (2022).

[45] OONI. [n. d.]. Global community measuring Internet censorship since 2012. https://ooni.org/. ([n. d.]). Accessed: 2023-2-15.

[46] Chiara Orsini, Alistair King, Danilo Giordano, Vasileios Giotsas, and Alberto Dainotti. 2016. BGPStream: A Software Framework for Live and Historical BGP Data Analysis. In *Proceedings of the 2016 Internet Measurement Conference (IMC '16)*. Association for Computing Machinery, New York, NY, USA, 429–444. https://doi.org/10.1145/2987443.2987482

[47] Ramakrishna Padmanabhan, Alberto Dainotti, Nima Fatemi, Arturo Filastò, Maria Xynou, and Simone Basso. 2019. Iran's nation-wide Internet blackout: Measurement data and technical observations. (11 2019). https://ooni.org/post/2019-iran-internet-blackout/

[48] Ramakrishna Padmanabhan, Arturo Filastò, Maria Xynou, Ram Sundara Raman, Kennedy Middleton, Mingwei Zhang, Doug Madory, Molly Roberts, and Alberto Dainotti. 2021. A multi-perspective view of Internet censorship in Myanmar. In *Proceedings of the ACM SIGCOMM 2021 Workshop on Free and Open Communications on the Internet*. 27–36.

[49] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nick Weaver, and Vern Paxson. 2017. Global measurement of DNS manipulation. In *26th USENIX Security Symposium*.

[50] Jonathan M Powell and Clayton L Thyne. 2011. Global instances of coups from 1950 to 2010: A new dataset. *Journal of Peace Research* 48, 2 (2011), 249–259.

[51] Lin Quan, John Heidemann, and Yuri Pradkin. 2013. Trinocular: Understanding Internet Reliability through Adaptive Probing. In *Proc. of ACM SIGCOMM (SIGCOMM '13)*. Association for Computing Machinery, New York, NY, USA, 255–266. https://doi.org/10.1145/2486001.2486017

[52] Ram Sundara Raman, Leonid Evdokimov, Eric Wurstrow, J. Alex Halderman, and Roya Ensafi. 2020. Investigating Large Scale HTTPS Interception in Kazakhstan. In *Proceedings of the ACM Internet Measurement Conference (IMC '20)*. Association for Computing Machinery, New York, NY, USA, 125–132. https://doi.org/10.1145/3419394.3423665

[53] RIPE. 2023. Routing Information Service (RIS). https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris. (2023). Accessed: 2023-2-8.

[54] Jan Rydzak, Moses Karanja, and Nicholas Opiyo. 2020. Dissent Does Not Die in Darkness: Network Shutdowns and Collective Action in African Countries. *International Journal of Communication* 14 (2020), 24.

[55] James Vincent. 2016. UN condemns internet access disruption as a human rights violation. https://www.theverge.com/2016/7/4/12092740/un-resolution-condemns-disrupting-internet-access. (2016).

[56] Nils B Weidmann and Espen Geelmuyden Rød. 2019. *The Internet and political protest in autocracies.* Oxford Studies in Digital Politics.

[57] Maria Xynou and Arturo Filastò. 2022. New blocks emerge in Russia amid war in Ukraine: An OONI network measurement analysis. https://ooni.org/post/2022-russia-blocks-amid-ru-ua-conflict/. (2022). Accessed: 2023-2-15.

[58] Maria Xynou, Moses Karanja, Berhan Taye, and Arturo Filasto. 2019. Resurgence of Internet Censorship in Ethiopia: Blocking of WhatsApp, Facebook, and African Arguments. (Aug 2019). https://ooni.org/post/resurgence-internet-censorship-ethiopia-2019/

[59] Jonathan L Zittrain, Robert Faris, Helmi Noman, Justin Clark, Casey Tilton, and Ryan Morrison-Westphal. 2017. The shifting landscape of global internet censorship. *Berkman Klein Center Research Publication* 2017-4 (2017), 17–38.