CORDON—A Taxonomy of Internet Censorship Resistance Strategies

Tariq Elahi and Ian Goldberg
Cheriton School of Computer Science
University of Waterloo,
Waterloo, ON, Canada
{*tariq.elahi, iang*}*@uwaterloo.ca*

***Abstract*—We present a taxonomy of Internet censorship resistance strategies and techniques extracted from analyzing proposed and implemented systems. We categorize the strategies into the following six types to form the CORDON taxonomy: *C*ollateral damage, where the damage caused by censorship would outweigh its benefits; *O*utside scope of influence, where the censor is powerless to act due to it having no control over entities or traffic; *R*ate limiting, where the censor's monitoring abilities are curtailed; *D*ecoupled communication, where bidirectional communications are asynchronous and asymmetric to take advantage of weaknesses in the censor's defences or hide the fact that the communications are related; *O*verwhelm, where the censor is deluged with large amounts of network traffic, paths and vectors to increase the cost and difficulty of effective censorship; and *N*o target, where the censor is unable to accurately detect and identify the people, infrastructure and network traffic to target. For each strategy, we identify the common supporting techniques used or proposed by resistance systems. We provide a detailed discussion of censors and their limitations, and outline the censor's decision-making process based on its utility, resources, and capabilities. We identify the censor's attack surfaces to provide context tying the censor capabilities and the resistance techniques together. We further identify future work needed to solve the fundamental problems facing all systems today: rendezvous protocols, bootstrapping without the need for client software, transports that stay ahead of the censor capabilities, and systems that scale better than the current batch. Applying CORDON to censorship resistance research will provide researchers a better understanding of the techniques and interactions that will help produce more effective and reliable censorship resistance solutions.**

***Keywords*-taxonomy; censorship resistance; censorship circumvention; systematization of knowledge**

I. Introduction

Censorship resistance, also known as censorship circumvention, is the response to the pervasive Internet censorship that exists today and continues to expand and adapt in step with technological advances and it is an active area of research interest. Either reactionary or preemptive, the aim is to mitigate the power of the censor. While there may be a plurality of motives, the basic goal is the same: free and open communication on the Internet. The focus of this work is on technological solutions for censorship resistance and we provide a systematization, as a taxonomy, of current censorship resistance knowledge. Our taxonomy, CORDON, will provide a better understanding of the general strategies, techniques and assumptions at play in previous works, give researchers guidance for future directions, and ultimately promote a freer and more open Internet today and in the future.

The Internet is a tool that impacts the lives of hundreds of millions of people around the world. It allows the fluid exchange of information and ideas from disparate corners of the world linking individuals together economically, socially and politically. The ease with which information can be disseminated through the Internet has been a boon to successfully creating social change to benefit residents of oppressive regimes. Recent history shows that the events of the Arab Spring were in part spurred by the ability of revolutionaries to organize and mobilize the population at large through the use of social networking tools such as Facebook and Twitter. Additionally, news of events being leaked from within—again through the Internet—engaged the rest of the world, bringing attention to the unfolding events, and applying pressure on the ruling elite.

Indeed, so successful is the Internet for dissemination and organization that oppressive regimes regularly curtail or outright censor it. These regimes are not alone as many governments, Internet service providers, corporations, and even households exert varying levels of control within their spheres of influence. There are legitimate reasons for controlling the spread of information such as privacy concerns, national security, corporate confidentiality, and public safety; however, there are also questionable reasons such as the chilling of speech, governmental largess, corporate misdeed, and subjugation, to name a few. This work does not attempt to sort through real world actors' motivations for censorship. We only assume that there exists Internet censorship, from the perspective of some entity, that ought not be. This can be a teenager stymied by parental controls or a whistleblower trying to collect evidence of wrongdoing.

Internet censorship is the, perhaps only universal, reaction of those in power to limit the impact of the Internet on their power base. An interesting effect of censorship is that it can present a view of the world that is aligned to particular ideals and can help promote the status quo and further control. Censorship has a long history in other mediums and it has been adapted to the Internet to deal with the technological, legal, economic, and political paradigms that have emerged.

Censorship controls can be technological or social in nature, complementing each other to provide the censor with the appropriate mixture of responses to suit their censorship policies—see §III-F for further discussion.

The battle between the two sides—Censor and Circumventor—is a never-ending struggle to gain the upper hand. As the censor learns how to be more effective, the circumventor learns how to overcome, which in turn teaches the censor. This cycle repeats until one side reaches the limits of the resources they are willing to invest, at which point an equilibrium is reached. The resulting balance, however, may later shift due to technological advances or changing resource limits.

The remainder of this paper is organized as follows: §II provides commentary on related efforts to systematize censorship resistance knowledge, and we present background on censorship in §III. §IV describes our proposed knowledge systematization, and we relate it to censorship and resistance technology in §V. We conclude with a discussion of future directions and areas of interest to the research community in §VI and make concluding remarks in §VII.

II. Related Work

Censorship resistance technology has steadily advanced and embraced many disparate techniques and fields. As the landscape becomes more complex, it is fruitful to organize it and provide context for reflection and as guidance for future efforts. Since the field is relatively young, efforts so far have been few and limited in scope and focus. We present here four previous works that classify censorship and censorship resistance strategies and techniques.

Köpsell and Hillig [26] present *blocking classification* from the point of view of the censor and its decision-making process, and also provide general *blocking resistance strategies*. While their main aim is to present a censorship resistance solution in a particular setting, their classification sheds light into the motivations and decision-making process of the censor and is valuable for this reason.

They classify blocking (or censorship) decision-making criteria according to the TCP/IP protocol layer of the communication and whether if it is based on the *content* or *circumstances* of said communication. They classify the *circumstances* of communication into addresses, timing, data transfer, and services. Examples of these include: source and destination addresses, send and receive times and connection duration, bandwidth and latency, and the protocols used. The *content*, where deep packet inspection is employed, is categorized by kind, properties, type and value. The kind of communication can be an object or a stream; its properties can include whether it is encrypted, compressed or other statistics. The type can tell us if it is an image, audio, or a binary, and the values are the byte patterns in the transmission that can be detected. It is feasible that as technology progresses, deep packet inspection will provide further identifiers that the censor could leverage.

They consider blocking resistance as concerned with two challenges. On the one hand is the *infrastructure* of the blocking resistance service and on the other is the *distribution* of the information about the blocking infrastructure. For the former they identify the "many access points" and the "all or nothing" resistance strategies. They give the example of deploying a large number of access points which the censor must counter completely or else have failed in censoring effectively. They identify the "out-of-band" strategy observing that the communication requirements for distributing information *about* the service are far lower than for traffic flowing through it. They also identify techniques for limiting the censor's ability to harvest information about the "many access points" but limit the discussion to their key-space hopping proposal.

We extend their work by providing a more thorough look at the strategies employed by a wider range of blocking resistance systems. We also explore the relationship between the capabilities of the censor and the attack surfaces present on the Internet.

Perng *et al.* [33] present a loose classification of publication censorship resistance schemes, whose goal is to ensure the availability of published information while denying the censor the ability to remove or alter it. They break down the resistance schemes into four categories: data replication, anonymous communication, server deniability, and data entanglement. Data replication prevents the censor from finding—or being able to delete if found—all copies of the target document. Anonymous communication is a building block of many systems and hides the user so that the censor can not find out who to target in the physical world. Server deniability protects hosts in the resistance system from being targeted by the authorities since they cannot control what is stored. Finally, data entanglement leverages the fact that there exist documents that the censor does not wish to suppress, which can then be "entangled" with documents that are not allowed by the censor. Removing one removes the other and the expectation is that the censor would not take this action.

While this is a limited classification, both in scope and sophistication, it helps to highlight basic strategies in the publication censorship resistance space.

In a current IETF-draft document, Barnes *et al.* [3] provide a discussion of filtering and its implications on transparent and correct Internet operation. They point out that filtering can occur at three places in the end-to-end communication channel: at the server side, at the client side, or in the middle. They argue that filtering in the middle is the most brittle solution, breaks security properties, and is not in keeping with the end-to-end ideals of Internet communication. However, because it is the cheapest method it is the most widely used, even though it is, according to them, effective only in constrained circumstances. On the other hand, filtering at the server and client side is a more natural fit to the Internet architecture; in particular, they expound that client-side filtering is, from an Internet architecture standpoint, the ideal choice. We contrast our work by observing that filtering of traffic is but one of many attack surfaces that form a more complete picture of the censorship resistance space. We also provide discussion of censorship strategies that apply to the traffic attack surface, and employ the filter techniques Barnes *et al.* address, and then go on to show how they are mitigated by the resistance strategies from our CORDON taxonomy in general and the systems that employ them in particular.

The closest related work to ours is also a draft, from 2010, where Leberknight *et al.* [27] provide a survey of censorship and its resistance from technological and political dimensions. They provide a taxonomy of resistance technologies, and discuss the design features that are critical for success. While similar in theme, our works diverge in several points. We are more concerned with the technological aspects of censorship and censorship resistance having analyzed proposed and existing systems while they focus in equal parts on the social and political as well. They only consider anonymity, content protection and content filtering as the challenges to overcome while we show that there are more attack surfaces for the censor to target. The techniques they discuss are limited to their identified challenges and the taxonomy they provide is similarly restricted to those aspects of censorship resistance. We provide a thorough treatment of the attack surfaces and censor capabilities, and provide more breadth and depth in terms of the censorship resistance techniques. Finally, since 2010 many novel systems have been developed that have added to the censorship resistance toolkit and thus our work provides a more contemporary account of the state of the art.

These works provide useful insights into the censorship resistance space, both from the censor and resistance system perspectives. We leverage these classifications as appropriate and augment them to produce a more robust and all-inclusive classification of censorship resistance.

III. Censorship and the Censor

The presence, or potential, of Internet censorship is what gives rise to censorship resistance. Therefore, it is important to describe censorship and the censor in detail to provide context and motivation for the discussions of censorship resistance systems that follow.

A. Censorship: A Definition

What is Internet censorship? This is a difficult question since the Internet is home to diverse entities with just as diverse needs, making it difficult to produce a definition that is universally applicable. Generally, the censorship resistance literature works around this problem by limiting the discussion to the specific requirements and target audience of the proposed system.

We similarly limit the scope of our definition to consider technological and social censorship and focus on technological solutions. We define it thus: *Internet censorship is the intentional suppression of information originating, flowing or stored on systems connected to the Internet where that information is relevant for decision making to some entity.* This definition allows for a censor that targets information for specific time periods, changes its stance after some time, or continues indefinitely. The entity in question can be an individual, a population, or a government upon whom the censor wishes to exert control or influence. The decision making may be as trivial as forming an unfavourable opinion or as extreme as organizing a rebellion to topple a government.

This definition provides context and motivation behind the censorship resistance systems and strategies we investigate.

B. Censor Attributes and Hierarchy

We categorize censors by their motivation, sphere of influence, and technical capabilities. While there are many types of censors operating under various conditions, in each instance the censor will have a motivation for its actions, but at the same time will be limited by its sphere of influence and capabilities. When responding to a particular scenario the censor must strike a balance between these constraints.

A censor's motives can span from the economic to the political and social. A business may want to restrict access to certain kinds of content that can be accessed on their premises due to fears of lost productivity, while parents may prefer to limit certain kinds of social networking sites their children take part in, and a government may want to limit speech for the sake of social order.

All censors have a limit to their power within and without their sphere of influence. These limits are relative and depend on the particular scenario. For example, while a government can exert a lot of control at the national level, parents can influence far more what occurs within the confines of their homes.

Similarly, all censors need to contend with their technical capabilities and limitations. Again these are relative; using the example above, while the government may have the most sophisticated filtering and blocking equipment, it may not compare to parents' watchful eyes on their children browsing the Internet. Conversely, the parent may not be able to effectively block certain resistance tools that the government is able to.

This suggests that a generalized censorship hierarchy is highly dependent on the scenario and the technological abilities of the censor. For concreteness we identify some examples of censors as reference points and show a graphical depiction of the relationship between resolution and scope

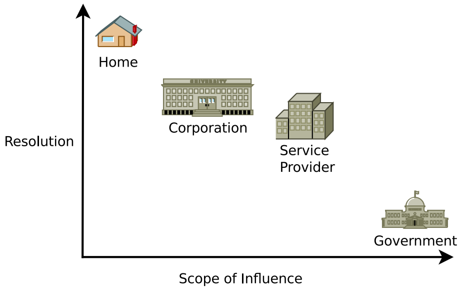

Figure 1. Mapping of examples on the space of resolution and scope of influence.

of influence in Figure 1. Here, resolution is the granularity with which a censor can observe the communications of the participants and be able to tell them apart while the scope of influence is the collection of systems, infrastructure, people and traffic that the censor has direct control or influence over.

- Government: This censor is motivated to stay in power either through good stewardship or oppression. It monitors internal traffic and that going through its borders for signs of unrest or other breaking of its rules. It can exert social and technological means to censor all traffic it can view. Depending on the country, it may further be willing to arrest citizens and invade their privacy. It has very little power outside its borders, unless it has international cooperation treaties in place.
- Service Provider: This censor is motivated by profit and continued existence. It can monitor all traffic that flows through its infrastructure. It can exert technological means to censor unwanted traffic, often supported or required by law. Cooperation with the Government may also take place to uphold the latter's policies. It has less legal power than the Government in scope but may have more fine-grained resolution and control.
- Corporation: This censor is motivated by productivity, efficiency and adherence to legal requirements. It can influence what occurs within its network and premises. It may not have the raw technical capabilities as the censors above but may have far more visibility into individual users' activities.
- Household: This censor is concerned with the moral and social aspects of the Internet and its effects in the home. Likely this censor lacks technical capabilities but it can still exert control through less technological methods of control such as parental filters and shoulder surfing. It is not likely able to influence much beyond the homestead.

C. Attack Surfaces

We now identify the *attack surfaces* used by censors to disrupt free and open communication on the Internet. These are physical or logical entities that can potentially be censored or attacked, and upon which the censor has visibility or influence. The following list provides an overview of these.

1) Traffic flows: Internet traffic can be monitored by the censor on its networks. The censor can regulate the traffic according to its censorship policies.

2) Infrastructure: The routers, servers and end hosts located on the censor's network or the Internet may be monitored for undesirable activity. While the censor may not have the same control over entities on the Internet as those on its own network it may be able to apply social, legal or technological pressure to achieve its desired result.

3) Clients: This is the client-side software that forms the conduit through which the user accesses the censorship resistance network, that manages the resistance mechanism, and through which censored content is published and delivered. The censor can target the client software distribution channels as well as the integrity of the software itself to erode the ability of the resistance system and users to connect together.

4) People: These are the publishers and consumers of censored information as well as the designers of resistance systems and the volunteers that facilitate access and publication. Along with technological tactics the censor can employ legal and extralegal methods in an effort to quash their activities.

5) Network-based views: Internetworking requires accurate views, such as routing and naming tables, of the network for correct [3], [21], and useful, operation. The censor can manipulate these views in order to limit the availability of, and access to, censored content. Depending on the situation, the effects may be felt by those on the censor's network, on the Internet at large [31], or both.

6) Censorship resistance system: The censor reacts to resistance efforts and causes the censorship resistance systems themselves to become a meta attack surface where all of the surfaces described above apply. The censor can attack the traffic, infrastructure, clients, people and network views that the censorship resistance system depends on. The censor can curtail the availability of censorship resistance systems specifically by making it difficult to establish first contact with them. They can do this by making it difficult to obtain information about them in general, interfere with specific discovery protocols, and finally interfere with the systems themselves.

The censor's motivation, sphere of influence and technical capabilities will dictate how, and to what extent, it leverages each of the attack surfaces above. For maximum efficacy, a savvy censor is likely to attack a number of the surfaces in combination.

D. Technical Capabilities

Censorship techniques fall into either *blocking* (sometimes called *filtering*) or *detection* categories and these cover the range of possible actions by the censor. It should be noted that none of the technological solutions for either is perfectly accurate or completely successful. It is possible that the censor targets honest communications by mistake, or unwittingly allows something it should not, so it must be careful when it employs these tools. As already mentioned, the applicability of these techniques depends on how well the censor is able to influence or observe the attack surface. We next present detection and blocking techniques as they relate to each of the attack surfaces we have identified.

1) Traffic flows: The censor can use packet header inspection to detect undesirable IP address and port destinations. Most modern firewalls allow this capability. The censor can go further by keeping track of the state of the links and performing deep packet inspection that analyzes the payload of packets for undesirable applications, strings and other patterns, or protocols. These methods incur additional overhead and thus the censor must be mindful of their limits.

To block at this level the censor can simply drop packets that it has detected as undesirable. In a more sophisticated approach the censor manipulates the traffic and expected protocol behaviour to break the normal flow of communication. For example the censor can try to interrupt SSL traffic by attempting to insert itself between the end points as a "man-in-the-middle". Similarly it can break protocols that expect the end points to adhere to agreed upon standards by injecting spurious traffic, as is the case where the Chinese firewall sends *RST* (reset) packets to both end points of the connection to break the communication flow [8]. This allows for more targeted blocking under certain circumstances where the censor is unable to safely block with a more blunt approach.

2) Infrastructure: The censor can use publicly available information to identify infrastructure, such as hosts and servers, that facilitate censorship resistance or the distribution of censored content. Often, where the network is "dark" or secret and known to members only, it can attempt to infiltrate the network by masquerading as an honest user or resource. In this way it can learn about its extent, operation and weaknesses. Internet service providers are a valuable resource in mapping Internet identities and addresses to their real-world counterparts. The censor can leverage legal pressure to extract this information.

To block infrastructure within its sphere of influence the censor can use legal, or extralegal, means to shut it down. Indeed, history shows that this has been successful [16] in the past. For infrastructure that is outside the censor's sphere of influence it can attempt legal measures but when this proves difficult it can attack the infrastructure through the network. An example of this is the distributed denial of service attack, which renders legitimate communication impossible due to the high volume of traffic directed at the targeted hosts and servers. Infiltration is not only useful for detection but also for launching attacks where the censor-controlled resources behave contrary to systems protocols to disrupt the communications or otherwise decrease the utility of the network.

3) Clients: The censor can compromise the client by installing monitoring software on it. This could be publicly announced and sanctioned by an authority, such as in the case of Green Dam Youth Escort [54] or TOM-Skype [42], or through covert means such as malware. This software performs content, keyword or destination filtering at the end host and may additionally report back to the authorities of the activities taking place on the end user's computer.

The censor can block at the client level by disallowing unapproved software from being installed on the operating system, disrupting functionality such as Internet searches, and displaying warnings to the user to dissuade them from attempting to seek or distribute censored content through the client or even to use it.

4) People: The censor may also identify people, using methods employed at the infrastructure level among others, who can be targeted to reduce the generation or consumption of censored material. Sometimes these people are public figures like Julian Assange [40] of Wikileaks [51] and at other times these people are well hidden as in the case of the members of the group called Anonymous.

The censor can remove these people from the equation by pursuing them through legal means, such as the U.S. government has done with Assange or through the threat of extralegal measures of a secret police, as is the case in Iran [23]. Note in the examples above that the censor has less control when the person is outside their boundaries [39].

5) Network View: Networking depends on nodes in the infrastructure having accurate information about the state and topology of the network. The censor can leverage this fact to corrupt the view and prevent clients from accessing the resistance network or censored content, and prevent the correct operation of the censorship resistance system. Of course, the censor can only act on those parts of this surface where it has influence. Usually, networking information is publicly distributed and includes routing, naming, and addressing. It is a simple matter for the censor to collect all resistance-related network information by inspecting the types of content usually hosted or retrieved. When this content is hidden from public view or encrypted the censor can again utilize infiltration techniques to identify the parts of the network to target. Once the censor has the "view" of the network that it wants to target it can proceed to blocking.

To block an accurate view of the network the censor provides erroneous information. Examples of this include DNS poisoning [3] and routing table changes that isolate traffic to known, trusted, and censored paths. There may be detrimental side effects [21], [31] due to the censor's activities on this surface since public network views are shared widely and expected to be accurate.

6) Censorship Resistance System: As a meta-layer, the discovery and the communication mechanisms of resistance systems can be targeted by the censor using all of the techniques that apply to the regular communications. Generally, it is expected that detection will be harder since the resistance tools are aware of common techniques and evade them where possible. However, sometimes there are quirks of the resistance system that distinguish it from regular traffic thus making it easier to detect; an example of this is the use of short lifetimes in the SSL [44] certificates that Tor routers used, which stood out from regular certificates that usually have longer lifetimes. Blocking can similarly be performed using many of the same techniques already described.

E. Capital, Tolerance and Utility

We assume that the censor is a rational actor, and it extracts utility from its censorship activities. Utility can be in terms of economic, social, and political outcomes, and there may be other aspects which can also be taken in to account.

The censor needs to ensure that its economic, social, or political capital is not negatively affected by its activities. Sometimes the will to enhance one facet can have a negative impact on the others; the censor must balance its desire for control against its tolerance to damage to its social, political and economic capital. Along with capital, this tolerance is an input for the censor's utility function. This multi-dimensional treatment provides a holistic view of the censor's limits and possible actions.

Indeed, Danezis and Anderson [9] provide support for this econometric treatment of censorship resistance that introduces utility as a metric for reasoning about censors' and circumventors' decision-making processes. In their model they demonstrate that by choosing certain resistance strategies a circumventor can degrade a censor's utility while maximizing their own. While their model does not take social and political implications into account and is restricted to trading off between two options, it is clear that treating the censor and circumventor as rational actors that want to maximize their utility is a reasonable approach and informs a deeper understanding of censorship and resistance dynamics.

F. Limitations: Technology and Policy

To round out this discussion we must point out that the censor has limits in the technical and policy dimensions. The hardware and software techniques the censor employs—as outlined above—have inherent scale, accuracy and resolution limits. While certain techniques are fruitful at smaller scales they become either impossible or too expensive to field at larger scales—scales that may be necessary for a particular attack surface scenario. Also, the censor's ability to perform targeted censorship is limited by a technology's resolution, meaning that it can not accurately discern one entity from another. A prime example of this is encrypted communication and the difficulty of differentiating legitimate traffic from censorable traffic at line speeds. Figure 2 provides a ternary plot of the censor's tradeoff between timeliness, resolution and accuracy for a fixed set of resources. The intersection of the dotted lines shows a Service Provider with an almost balanced trade-off between the variables. Also shown are a Home censor who may require high resolution and accuracy but is not concerned with timeliness; a Corporate censor that has less resolution and accuracy but higher timeliness, and the Government censor who may be concerned with bulk censorship that is accurate and timely but is not interested in the details of those it censors. With the addition of scope to this analysis the censor can chart out the limitations and strengths of different censorship scenarios in many dimensions.

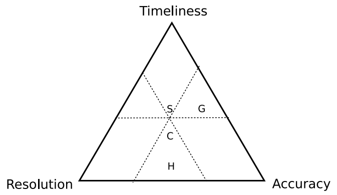

Figure 2. The points on the triangle denote the distribution of resources among the three variables. G, S, C, H denote Government, Service Provider, Corporate Entity, and Home respectively.

Often one aspect, such as economic policy limitations, supersedes all other considerations and the censor is then bound to cater to it, even if it means relaxing its censorship posture in other aspects. For this reason, censorship resistance systems often leverage the censor's tolerance with the assumption that the censor will be mindful of the utility of censorship. An example of this where the censor allows encrypted traffic to flow unhindered because it is necessary for e-commerce applications, and necessary for its economy, with the knowledge that encrypted resistance traffic is also getting through. Blocking this traffic is within the abilities of the censor but in this instance the economics of the situation preclude this option.

The circumventor has similar restrictions in terms of capabilities and resources and can use the same tools to perform a similar analysis in order to gauge the effectiveness of their resistance efforts. Of course nothing stops both sides from performing this analysis on each other and using it to tune their efforts. The censor must 1) develop its technological capabilities such that 2) they can be deployed against attack surfaces at the scale and resolution required while 3) balancing the harm to its capital. The circumventor must leverage 1) the technology and 2) policy limitations faced by the censor while also 3) balancing the costs associated with these efforts.

IV. Censorship Resistance

We now identify general resistance strategies and supporting techniques abstracted from our analysis of systems that

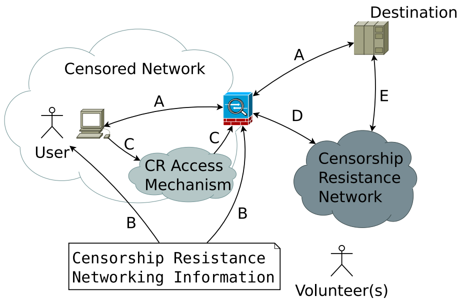
Destination
Censored Network
A
A
E
D
C
C
User
CR Access Mechanism
Censorship Resistance Network
B
B
Censorship Resistance Networking Information
Volunteer(s)

Figure 3. Censorship Resistance Channel Data Flows. Arrows A-E depict the channels of information flow for normal Internet traffic (A) and that for resistance traffic (B-E).

have been proposed in the literature, been implemented, or are in current use. It is our goal, through the application of our taxonomy, to discover gaps in the strategy and techniques space where possible. We will discuss the systems we analyzed and how the strategies apply in each and the techniques they employ. We will also discuss the relevance of these classes on the attack surfaces and technical capabilities of the censor, and the implications on the utility of the censor and circumventor.

A. *Overview*

We first provide an overview of censorship resistance from an abstract data channel perspective. Figure 3 provides a data flow diagram depicting these channels.

When there is no censorship, a user's data flows unhindered through the Internet gateway/firewall and onward to its destination as shown by the flows labelled *A*. This is the usual manner in which Internet traffic is meant to flow. In the presence of censorship, however, these flows are blocked, and alternative flows must be employed.

To do so, the user acquires client software and installs it on her computer. This software becomes the channel for all censorship resistance data flows. The software gathers information about the resistance network. It could be the case that this information and client software were delivered to the user surreptitiously via email through the censor's firewall, or were hand delivered by a volunteer. This flow of software and resistance network information is depicted by the data flows labelled *B*.

There are two types of censorship resistance a system can provide: publication and access. Publication censorship resistance prevents a censor from restricting the storage and availability of censored information. Access censorship resistance provides a means to access censored information, or to a publication censorship resistance system, when the censor has taken steps to prevent access to either. We will see how this affects data flow and channel use shortly.

To communicate, the client connects to the resistance network through an access mechanism that circumvents the censor's controls. For example, this mechanism can be as simple as connecting directly to a proxy server that is not blocked or as elaborate as a multistage protocol that requires the user to solve puzzles or compute mathematical functions to learn information about access points. Systems solely intended for publication assume that access to the system is not itself censored, or that a suitable censorship-resistant access system is in place. Once the client can establish a communication channel with the resistance system the interactions proceed according to the resistance scheme employed. These interactions are depicted by the flows labelled *C* and *D*.

It is important to note that the destination, or objective of the user, can be internal or external to the resistance system. Generally, destinations are internal to publication systems while most access systems are independent of destinations. Data flow *E* depicts the latter case.

We can further map the attack surfaces to show how they interact with the data flows and entities. The users and volunteers are both within the people surface; the data flows *A–E* are within the traffic surface; the censorship resistance network (and sometimes destination) as well as the client software are part of the infrastructure surfaces; and finally at the meta-level the client software, resistance network, network information, access mechanism are volunteers are all part of the censorship resistance surface.

B. *General Classes of Resistance Strategies*

We now present the general classes of resistance strategies and techniques. These strategies and techniques have been systematized into the CORDON taxonomy of censorship resistance that is represented in Figure 4. For each strategy, we identify the common techniques that support it that have been used or proposed by resistance systems.

1) Collateral Damage: One of the more powerful strategies is to cause collateral damage to the censor's capital as a result of its censorship activities. This damage can be to economic prosperity, social harmony, or political control. The key is to accurately judge the censor's technological and policy limitations and utility and then to produce situations where collateral damage to its capital is unavoidable.

In each of the descriptions below the censorship resistance technique assumes that the censor is unable to effectively tell censored and allowed content apart. In all instances if the service, information, or path became inaccessible there would be a detrimental effect on the censor's capital. It also assumes that the censor does not have viable alternatives—which is not always the case [30]—for any of the services, paths, or information that the resistance system leverages.

A popular technique is to hide among innocents. This resistance technique embeds its information with other, useful, information. The censor, being unable to sort them apart, is bound to either remove access to all the information or allow it.

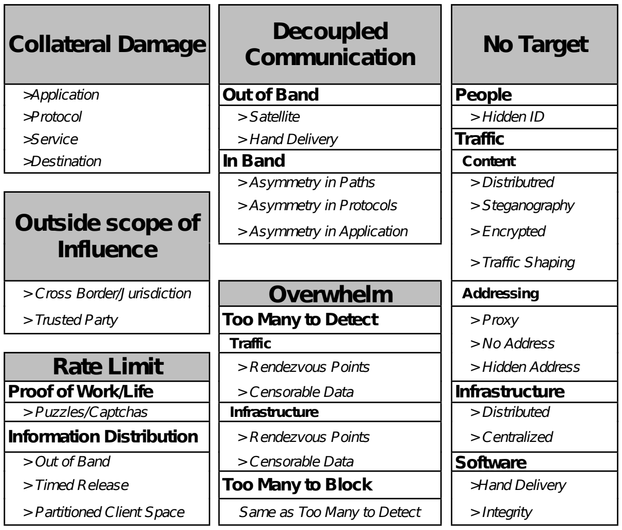

Figure 4. The CORDON taxonomy of censorship resistance strategies. Along with each of the six strategies, we list a number of techniques that support it.

Another technique is to hide on the path to resources valued by the censor. The resistance system utilizes these paths that are used to access regular, and useful destinations, but uses it to transmit censored content. The censor, being unable to block the path due to its importance, must either allow all communications on it or cut itself off from it.

Finally, the resistance system can make use of an existing and useful application or service the censor is loath to cut itself from due to the impact it would have on its capital [41]. Services such as free email, *e.g.* Gmail [49], and cloud computing platforms, *e.g.* Amazon Web Services [24], have been leveraged by resistance systems for this reason.

2) Outside Scope of Influence: An inherent weakness of any censor is the inability to control or influence activities and entities outside its sphere of control; see §III for concrete instances of this phenomena.

Resistance systems leverage cloud-based services, nodes and other entities across multiple political borders, and entities that are trusted not to be coerced by the censor as resistance building blocks. The key to effective resistance is understanding the limits of the censor's influence and leveraging entities that fall outside it.

3) Rate Limiting: A weakness of any resistance system is that users should be easily able to find information about them and how to access them. However, it is difficult to tell a censor and an honest user apart, which makes hiding information from the censor challenging. The censor may be able to effectively and efficiently harvest information about the resistance system and then neutralize it. The harvesting may be so effective, *i.e.* timely and accurate, that it renders even the overwhelming techniques (see below) redundant. To combat this, many resistance systems employ schemes to limit the rate of information harvesting. They must balance the ease of use for users and the need to restrict the censor's harvesting effectiveness while allowing an acceptable level of censorship resistance to occur.

Puzzles, such as captchas and computational tasks, have been proposed [14], [29] to slow automated harvesting and which may cause the censor to employ human beings and thus raise their cost of effort.

Information distribution techniques that restrict information flow to subsets of recipients or time slices, such as keyspace hopping [14] or Tor's bridge distribution scheme [10], increase the overhead and time requirements respectively for the censor to learn enough about the network to effectively neutralize it.

4) Decoupled Communication: Upstream and downstream Internet communication requirements are sometimes asymmetrical and lend themselves to the leveraging of atypical communication protocols and communication mediums.

Keeping Kerckhoffs' Principle in mind, resistance systems should not depend on the secrecy of the decoupling mechanisms used. However, in practice, there is often a window of opportunity where the censor has not yet anticipated or is unable to effectively detect and block channels of communication that are utilized in a seemingly innocent and dissociated manner from the usual censorship targets. Leveraging the censor's blind spots about the nature of the bidirectional resistance traffic and by changing the signature of the bidirectional communication the circumventor can thwart the censor.

Two independent aspects are key: 1) The censor does not suspect that the individual communication is for censorship resistance and 2) that it is difficult to correlate the related asynchronous resistance communications streams. These streams can have characteristics, such as low throughput [13], [22], [50], that would make their universal use unacceptable but which still provide effective communication for specific use cases.

Simultaneous use of multiple protocols and transport mediums is an extra layer of confounding requiring the censor to monitor many fronts at once and correlate over a larger set of information.

5) Overwhelm: Leveraging the fact that the censor does not enjoy limitless resources and must react within a reasonable time window a powerful strategy is to overwhelm the censor through sheer numbers. The aim is that the censor will not be able to deal with every single vector and that even a single unblocked one will allow communication to occur.

There are two basic points of leverage for the circumventor. The first is the limits on the ability of the censor to detect censorable activity on surfaces at a given scale in an amount of time that does not negatively impact the censor's capital, when the censor is limited in the amount of resources that it is capable of mustering. The second leverages the fact that the censor is unable to effectively block targets at a certain scale or resolution due to technological or policy considerations.

An example of the former is the deployment of a massive number of proxies that the censor then has to discover, which takes time especially when rate limiting is employed, during which the resistance is effective. By extension, as an example of the latter, blocking becomes a policy issue if the proxies are usually honest nodes on the network and blocking them would cause a large part of the Internet to become unreachable and thus harm the censor's capital.

6) No Target: The circumventor attempts to fool the censor into thinking that resistance activities on the various attack surfaces it is monitoring are legitimate and hence not targets. It can do this by making its traffic look like legitimate traffic, or making the protocol match allowed protocols, or hiding its traffic within other traffic, or making it seem as if the communication is between allowed endpoints.

One way of making resistance traffic look legitimate is thorough the use of traffic shaping to match the characteristics of traffic that is allowed by the censor. The circumventor normally attempts to match traffic the censor deems important enough not to disrupt; otherwise, once the resistance technique is discovered, the censor can simply block the previously legitimate traffic. Often, there are windows of opportunity, bearing in mind Kerckhoffs' Principle as noted earlier, where the censor does not have a fingerprint to match against, or has old information, and the circumventor ensures that it does not match patterns of its past behaviour.

Making the communication look like a specific application or protocol is a slightly more sophisticated form of traffic shaping where the circumventor tries to use an application important to the censor as its cover. While the same conditions as purely traffic-based cover apply, the difference is that this cover is more believable when the application is in wide use around the world. For example, the TLS protocol, which is the linchpin of e-commerce communication, is widely used and has been leveraged by resistance systems.

Steganography, which is the technique of hiding covert information within innocuous communication, is also a viable resistance technique. The fitness of this technique for the purposes of censorship resistance is still under debate since the data rates which it can support are not as high as other comparable techniques and the rates of its detection are also not always encouraging. Still, where limited information transfer is needed then steganography can be useful.

Finally, circumventors can make it seem as if their communications occur between allowed endpoints. The primary example of this is the use of proxies. There are many types of proxies and indeed most resistance systems utilize proxies in some form; more details of the different types follow in §V. The inherent problem, and one that is of great importance, is of distributing information about the proxies to honest clients without the censor also finding out. The literature has instances of techniques employing trust-based [35], rate-limiting [10], [28], [29], and side-channel [10], [13] distribution to name a few.

Additionally, the resistance system protects the people, traffic, infrastructure and software by making them difficult to find and target. The censor cannot block an unaddressable or unidentifiable entity on the network. Since detection and blocking on IP and port combinations is cheap and readily available, resistance systems attempt to hide this information. This can happen when the entity is a node that has no address such as a network router, or when the node addresses are unknown, for example when onion routing hides the source or destination of a flow from certain observers, or when the circumventor is anonymous. Resistance systems that leverage routers [19], [25], [55] are chief examples of the former case. As long as the censor is unable to compromise the router or route traffic away from it [34] they are unable to effectively deal with the systems that use routers as the point of resistance. The censor is similarly unable to block nodes it is unable to learn the addresses of [11], [49]. Nodes can hide their addresses using a gateway [29], routing schemes [7], or behind a cloud service [24].

Targeting non-network entities such as people can be difficult if they are well hidden [38] using anonymous communication and side channel means or when the resistance system is autonomous and self organizing [7], [35] and out of the control of any individual person.

C. Attack Surfaces Resistance

We now relate these censorship resistance classes to the attack surfaces, both for regular communication and that of censorship resistance systems. The robustness and efficacy of the classes will depend on the censor's abilities as were discussed in §III.

1) People: Using anonymizing techniques, consumers and publishers of censored content provide *no targets* to the authorities. Anonymization can be used by the volunteers, makers, and administrators of the resistance system who can also be *outside scope of influence* of the censor and sometimes have *large numbers*.

2) Traffic: Presenting *no targets* and *decoupled communication* are the main mitigation strategies for protecting censored traffic. Encryption [11] and traffic shaping [32] are the most widely used techniques to achieve this. Recent years have seen the rise of asymmetric and asynchronous communication as a means to decouple end point communication.

While not an electronic means of transmission but in keeping with the theme of decoupled communication, the physical transportation of client software and censorship resistance network information is also a viable technique which almost all systems [10] can leverage.

3) Infrastructure: The routers, servers and end hosts of censored content can be shielded by causing the censor *collateral damage* through leveraging popular services and destinations, presenting *no target* by hiding the locations and identities of nodes and servers, being *outside scope of influence* by deploying across many judicial and national boundaries, and relying on *overwhelming* numbers of access points.

4) Network Views: To protect and ensure truthful reporting of network views, resistance systems can retrieve information using *decoupled* channels via a trusted party *outside the scope* of the censor's influence, employ *rate limiting*, and present *no target*.

By receiving accurate network view information through decoupled channels that are outside the censor's scope of influence, such as hand delivery by a trusted resistance volunteer, the user can be assured that their client has a complete view of the network and provides the best user experience. Where the censor blocks these channels or infiltrates the user base or resource pool, the network and its view can be preserved by releasing network information slowly over time or breaking it into non-overlapping subsets—in order to prevent any entity from learning about the whole network—and distributing it among users.

A big step in mitigating the network view surface is to obviate the need to maintain a view at all. Removing any addressing, routing, and naming information from censored communications renders the censor impotent in this regard. While this is a positive paradigm shift, Schuchard *et al.* [34] show that it is very difficult to achieve this in practice.

5) Clients: To protect the client surface the same strategies for receiving network view information can be employed. *Decoupled* communication channels, *e.g.* hand delivery of client software, is a viable strategy. To mitigate the threat of eavesdropping by external entities, the client can attempt to hide its activity using techniques to present *no target*, such randomizing [43] its activity.

Currently, no existing system hides the fact that the end user is using censorship resistance software from an observer who can monitor the activities performed on the user's own computer. This is a weakness that needs to be addressed as it also impacts the people surface.

V. Censorship Resistance Systems

We now apply the CORDON taxonomy to existing censorship resistance systems. We group systems by their basic functions: publication or access (or both), as defined in §IV-A, or facilitation. As it happens, research interest has also trended from publication to access censorship resistance so our treatment will be mostly chronological and strategies will be introduced in order of first use. Table I summarizes the techniques and strategies used by each of the resistance systems mentioned above with the year they were introduced and the basic resistance functionality they provide. For context, refer to Figure 4, which categorizes techniques into the strategies they support, and §IV-B, which provides supporting details.

A. Publication

In 1996, Anderson's Eternity Service [1] drew attention to Internet censorship resistance research on the publication side. His proposed system was the first to leverage distribution of data across a large number of servers which are deployed in diverse jurisdictions. Back's Usenet Eternity [2] was a 1997 implementation of this proposal.

Utilizing the Remailer concept based on Chaum's ideas [6], Goldberg and Wagner's Rewebber proposal [17] utilizes multi-hop routing for encrypted publication on servers whose identities are hidden behind a pseudonymous name space. This provides no target for the censor to block since it is difficult to track the servers down and even if that were done it is impossible for the server to know what content it hosts due to encryption, it is therefore impossible to target specific content. Similarly, Publius [48] and Tangler [47] both use this same trick of encrypting all stored content for plausible server deniability properties. The difference is that in Publius the encryption key is shared among servers using a k *out of* n threshold secret scheme so that some lost, *i.e.* censored, shares are acceptable. Tangler on the other hand introduces the concept of collateral damage by enforcing, through encryption, that any single document depends on other documents and to remove it would mean the removal of these other documents also. Where the censor has vested interests in certain documents remaining available, this technique provides good protection.

Freenet [7] introduced the use of peer-to-peer (P2P) overlay networking to provide censorship resistant storage of documents. Search and retrieval operations cause redundant copies of documents to be made thus ensuring their long-term availability. The decentralized topology ensures that the censor can not target the true location(s) of documents and also leverages the fact that most servers will be beyond the censor's influence. Building on top of this, Serjantov [36] introduces anonymous access and the concept of "forwarders" as a defence behind which the "storers" of documents can hide and thus evade the censor.

Although not P2P in nature, Tor Hidden Services [11] also leverage anonymous access to hide the true location and identity of servers from not only the users but also from the network nodes.

While most research remains on paper it is interesting to look at which systems become implemented and remain relevant. Tor Hidden Services and Freenet are in operation today although it is unclear just how popular these services are—a side effect of the strong anonymity and privacy guarantees.

B. Access

While publication censorship resistance concentrated on the infrastructure and people attack surfaces, access cen-

Table I
CENSORSHIP RESISTANCE TECHNIQUES LEVERAGED BY VARIOUS SYSTEMS, AS DESCRIBED IN SECTION IV AND FIGURE 4. P, A, F STAND FOR PUBLISHING, ACCESS AND FACILITATION RESPECTIVELY.

System	Year	Collateral Damage	Outside Scope	Rate Limiting	Decoupled	Overwhelm	No Target	Function
Eternity [1]	1996		Jurisdiction			Rendezvous		P
TAZ-Rewebber [17]	1997		Border				Encrypted, Hide Address	P
Web MIXes [4]	2000		Border				Proxy, Encrypted, Hide Address	A
Freenet [7]	2000						Dist., Hide ID/Address	P
Publius [48]	2000						Dist.	P
TriangleBoy [20]	2000		Jurisdiction	Out of Band	Asym. Paths	Rendezvous	Encrypted	A
Tangler [47]	2001	Service					Encrypted, Dist.	P
Serjantov [36]	2002						Dist., Hide Address	P
Infranet [13]	2002	Service			Asym. Protocol	Rendezvous, Data	Steg.	A
Untrusted Messenger [14]	2003			Partition		Rendezvous		F
Tor-Hidden Services [11]	2004						Hide Address	P
Köpsell and Hillig [26]	2004		Border	Puzzles			Proxy, Hide Address	A
Clayton *et al.* [8]	2006						Traffic Shaping	F
Tor Bridges [10]	2006		Border	Out of Band, Timed			Hide Address	A
Dust [53]	2010						Encrypted, Traffic Shaping	A
Collage [5]	2010	Service			Asym. Application	Rendezvous, Data	Steg., Dist.	P+A
Proximax [28]	2011			Partition			Dist.	F
COR [24]	2011	Service	Jurisdiction					F
Telex [55]	2011	Destination					No Address	A
Decoy Routing [25]	2011	Destination					No Address	A
Cirripede [19]	2011	Destination					No Address	A
Freewave [18]	2012	Application					Hide ID/Address	A
MIAB [22]	2012	Service			Asym. Application		Steg., Dist.	P+A
FlashProxy [15]	2012	Destination				Rendezvous	Proxy, Hide Address	A
DEFIANCE [29]	2012	Destination	Jurisdiction	Timed Release		Rendezvous	Hide Address	A
Obfsproxy [45]	2012	Protocol					Traffic Shaping	A
StegoTorus [50]	2012	Protocol					Steg.	A
SkypeMorph [32]	2012	Application					Encrypted, Traffic Shaping	A
CensorSpoofer [49]	2012	Application, Destination			Asym. Application		Proxy, Traffic Shaping	A
Unblock [35]	2012			Partition			Dist., Hide Address	A

sorship resistance focuses on the traffic and network view surfaces.

Web MIXes [4] utilizes cascades—routers in predetermined and fixed chains—and client- and server-side proxies to provide unlinkable traffic between the user and host. This ensures that the censor does not learn the identities of those involved, to then target. However—and this is in common with many other systems—the addresses of the access points are well known and the censor can simply block all communications with those addresses. As a countermeasure, Köpsell and Hillig [26] extend the proxy design by having a large number of proxies behave as network access, or rendezvous, points for the client. To limit the rate of effective harvesting of proxy addresses by the censor, a puzzle-based challenge is presented when requesting a proxy address. This requires the censor to expend effort to obtain proxy addresses, and thus makes automated, or human-driven, harvesting more expensive and time consuming.

TriangleBoy [20] and FlashProxy [15] also utilize large numbers of proxies, albeit with important differences. TriangleBoy proxies only relay the requests from the client, introducing a decoupled communication scheme where requests and responses flow over different network paths and network protocols. This decoupled behaviour is leveraged by other systems as we shall see below. The proxies in FlashProxy exist within a user's browser—when the user visits a supporting website, Javascript proxy code runs on the user's computer. This proxy exists for only as long as the client is viewing the website. This transient nature makes blocking inefficient since blocking of short-lived proxies requires that the censor either update the list frequently or suffer from overblocking when those proxies cease to exist and their addresses potentially become used for legitimate needs.

Tor [11] is a popular anonymous communication system that has increasingly been used for censorship resistance since it allows access to censored content. Since Tor's rendezvous points, or routers, are public knowledge the censor can simply block access to these addresses from its users. Tor bridges [10] were introduced in response, to break through this form of blocking. Tor bridges are just like regular Tor routers except their addresses are only distributed slowly—by word of mouth and email, for example. A number of addresses are held back from distribution as reserves to counteract efficient address harvesting by the censor. Clients connect to bridges which then connect to the Tor network as usual to circumvent the censor's blockade. The censor tries to counteract this by scrutinizing the traffic and network view surfaces for Tor-like traffic and behaviours.

Tor pluggable transports [46] and related systems protect the traffic attack surface by making it difficult for the censor to detect Tor communications. Obfsproxy [45] and StegoTorus [50] both transform Tor traffic, which is encrypted and therefore something the censor looks for among other patterns, to mimic other types of traffic, *e.g.* HTTP. StegoTorus, which is part of the DEFIANCE framework described below, further uses steganography when transforming the Tor traffic, although this does reduce the bandwidth of the system.

DEFIANCE [29] is a framework that addresses Tor bridge detection and address harvesting. The system provides security by layers. There is a gateway that gives access to a restricted set of bridge addresses which are usable during a limited time period. Clients are required to perform computations and follow a specific interaction in order to receive bridge information and maintain communications. These techniques hinder the ability of the censor to compile an accurate and complete list of bridge addresses.

The censor can use the quirks of Tor's protocols to fingerprint it. Wiley introduces Dust [53] to remove the fingerprint of Tor's plaintext TLS handshake traffic by encrypting it. It overcomes the censor's ability to fingerprint string patterns, packet lengths, and timing information of a Tor communication stream. It also obviates the need for traffic-level steganography. Wiley also provides an automated means of testing blocking-resistant protocols using Bayesian classification to further help test and enhance them. [52]

Transforming traffic to look like another protocol, such as HTTP, generally has only limited success, as the cover traffic does not appear natural. The censor may be able to distinguish the artificial Tor-disguised-as-HTTP traffic from regular HTTP traffic, and block it. Similarly, sending streams of encrypted bytes not corresponding to any known Internet protocol is also straightforward to detect, and may attract the attention of the censor. However, combining these techniques can yield a powerful tool: disguise communication as an existing popular encrypted protocol. Three recent systems, SkypeMorph [32], Freewave [18] and CensorSpoofer [49], leverage Skype [37] communication protocols for this reason. SkypeMorph shapes the traffic of Tor communications to look like that of a Skype video call while Freewave converts the data into sound signals that are transmitted over a Skype voice call. CensorSpoofer utilizes decoupled communication channels, akin to TriangleBoy. Clients request content, protected by steganography, through out-of-band means such as email or instant messages that the censor is not likely to block. At the same time the client establishes a VoIP connection to some unblocked host as a dummy destination. The CensorSpoofer proxy sends the responses to the client over the VoIP channel and rewrites the source address to be that of the dummy host in order to fool the censor. All of these schemes rely on the fact that the censor does not have any influence on the Skype client software and its operations and cannot corrupt the client's view of the Skype network.

Infranet [13] also utilizes steganography to hide information in images that are posted on popular image-hosting websites as well as a "knocking" protocol to access the images. These images are retrieved by performing a series of *GET* "knocks" by the client that matches a predetermined pattern. The addresses of the image hosts are carefully distributed in order to avoid harvesting and blocking. The aim is that the censor remain unaware that censorship resistance is afoot.

Recent systems have introduced a method of resistance that obviates the need for learning the rendezvous addresses of the resistance system. Decoy Routing [25], Telex [55], and Cirripede [19] leverage details of encrypted Internet protocol traffic to send signals to routers en route to allowed ("overt") destinations to divert the flow of the traffic to censored ("covert") destinations. The only conditions are that 1) the router is on the path to overt destinations that would cause large collateral damage to the censor if they became unreachable and 2) there be no alternative routes that avoid the router. The main differences between the systems are in the encryption schemes used. Decoy routing is based on symmetric encryption while both Telex and Cirripede are based on asymmetric encryption. Decoy routing sends a sentinel, a symmetric key the router and user share, along with the hello message, whereas Telex tags a TLS random nonce using public key steganography and Cirripede does something similar but with the initial sequence number found in the TCP header. The trick is that the censor is unable to tell normal encrypted traffic from decoy routed traffic but the router on the path is able to do so with little effort.

Unblock [35] leverages existing social trust links within P2P overlay networks such as Freenet [7]. If the users of the network only connect with known friends who will not inform the censor then the social graph, and network, will be resilient to infiltration. Due to information locality—*i.e.* only information about neighbouring nodes is exposed—harvesting will be restrained and thus effective blocking will be difficult.

C. *Access and Publication*

Systems that provide censorship resistance for both access and publication are few. Collage [5], which extends the basic ideas of Infranet [13], and Message in a Bottle [22] (MIAB) are two such systems. Due to their dual functionality they utilize many strategies at once. Both Collage and MIAB leverage evasion, large numbers, collateral damage and decoupled communication as the primary strategies. In Collage, steganography is used to hide messages in images posted on popular picture-hosting websites and thus evade the censor's filters. MIAB also utilizes image steganography and further enhances this by utilizing blog posts instead of just posting the image on a popular hosting service. The advantage of both systems is that no extra infrastructure is required on top of the existing blog and image-hosting services, and the censor is hesitant, due to economic, social and political reasons, to restrict access to a useful and popular service. In both, rendezvous is accomplished by ensuring the communicating parties are subscribed to likely and frequently visited, but not easy to identify as censorship-circumventing, services and accounts.

Unfortunately both these systems suffer from large communication overhead and are not useful for usual Internet activities such as browsing and bulk transfers. They are most suitable as bootstrapping mechanisms to establish communication channels with higher bandwidth.

D. Facilitation

Some systems only address specific challenges in censorship resistance and could be utilized to address the shortcomings of the systems above.

Network information dissemination is a common challenge in censorship resistance systems, as it is hard to know if we can trust that the recipient is a user and not the censor. A censor can masquerade as an honest user to collect information about all proxies, and block access to them all, cutting censored users off from the outside world. Both Feamster *et al.* [14] and Proximax [28] provide solutions to mitigate this issue.

Feamster *et al.* utilize *key space hopping* as a means of partitioning the proxy address space in a way that is dependent on attributes of the user, in this case their IP address. Thus, the adversary needs to masquerade as many users in order to collect the values from the entire address space. To further hinder the censor, proof of work and life schemes, such as solving puzzles or captchas, are employed. They introduce a hierarchy of a few trusted proxies and many untrusted messenger nodes as a means of mitigating proxy exposure. On the other hand, Proximax assigns trust by measuring the safety of distribution channels. Each channel (a user) is assigned a subset of proxy addresses, which are polled for availability. Each channel is expected to redistribute the available addresses to other as yet untrusted users. If the proxies remain active then that channel is to be trusted further, but if they become unavailable, or unreachable, then the channel is not to be trusted and further proxy information is not distributed through it.

While trust is a useful, yet difficult to measure, metric, some systems depend on diversity to provide safety. Cloud-based onion routing [24] introduces cloud-based hosting as a means to further enhance the diversity of the Tor network. By involving a large number of service and infrastructure providers, located in diverse jurisdictions, the level of diversity will increase and thus is will be unlikely that the censor can pressure all the parties and also equally unlikely that they would cut themselves off from so many Internet destinations and services.

Sometimes the shortcomings are in the censorship systems themselves. Clayton *et al.* observe that, circa 2006, the Chinese censorship system utilized an easy-to-overcome technique to block traffic to undesirable websites. TCP reset packets were being used to break connections between endpoints and the remedy was to simply ignore these. While the utility may be limited it is important to realize that censorship resistance systems are not infallible and that constant and close scrutiny may provide other effective resistance tricks.

VI. Future

We now discuss the implications of our findings on the future of censorship resistance research. We have identified five challenging areas that require research attention; these areas involve all the attack surfaces we have identified in this work.

The dependence on client software is a stumbling block for scalable and wide-reaching resistance solutions. Currently, most proposals expect that a trusted out-of-band mechanism exists—sometimes by hand—to transport client software to where it is needed. This means that the rate and extent of distribution is limited to the couriers' opportunistic personal meetings and groups of contacts. It excludes those users who do not have regular contact with volunteers or frequent the same places.

Much more progress needs to occur to secure the traffic attack surface. The censor's tools become more sophisticated with the passage of time and it becomes more difficult to counteract them [12]. Perhaps it may be fruitful to identify and extend the traffic attack surface into areas where the censor is weak, such as out-of-band communication channels that have performance and security characteristics similar to the Internet but do not rely on using infrastructure owned or influenced by the censor.

Dependence on network views is a weakness that should be overcome because most often the censor controls this information and is apt to manipulate it to its advantage without too great a loss of functionality. Although we have seen the emergence of decoy routing [19], [25], [55] the censor already has possible counter measures [34]. Future systems should operate under the assumption that the network view is untrustworthy and should aim to provide useful resistance functionality in that environment.

Establishing first contact with the resistance system is very challenging. It is difficult because it is tricky to establish trust with strangers when the censor is among them. Barring unscalable proposals utilizing webs of trust, no proposal has so far overcome this shortcoming. Also, the resistance systems have to become more scalable if they are to be viable in the future. We take examples of existing systems, such as Tor, that are struggling to maintain levels of performance as more users come on board. In a hypothetical world where all network traffic is censorship resistant, existing systems would not fare well.

Finally, we posit an ultimate censorship scenario where an all powerful censor has control of all servers, traffic channels, and has far reaching influence. Indeed this censor is not so far fetched; China has steadily been copying popular parts of the Internet to 1) obviate the need for traffic to leave the country and 2) to reduce the impact of collateral damage of blocking foreign services; *e.g.* they have their own versions of Facebook and Twitter (Weibo), and Google (Baidu), and have been restricting their citizens' access to more and more of the Internet. By considering this extreme scenario we may learn how to provide censorship resistance in this extreme environment if the time should come when Chinese netizens are cut off from the Internet for good. The experience and knowledge gained from this activity can help ensure free and open communication on the Internet at large.

VII. Conclusion

Censorship resistance research is on the rise, as evidenced by the large research output in the last couple of years. We presented the CORDON censorship resistance taxonomy that categorizes this research into abstract resistance strategies and their supporting techniques. We analyzed the censor and the parameters of its decision-making process in terms of its utility, capabilities, and economic, social, and political capital. We presented an inherent technological tradeoff between accuracy, resolution, and timeliness that the censor need contend with when overcoming censorship resistance. We identified censorship attack surfaces and used them as a common reference point to discuss the censor's and circumventor's technical capabilities and strategies. Finally, we provided direction to future research efforts to overcome certain fundamental problems facing resistance systems today.

VIII. Acknowledgements

We thank NSERC, ORF and The Tor Project, Inc. for funding this research. We would also like to thank Ryan Henry for helping us formulate the definition of censorship that we use and Hooman Mohajeri Moghaddam for editorial feedback.

References

[1] R. Anderson, "The Eternity Service," in *Proceedings of Pragocrypt '96*, 1996.

[2] A. Back, "Usenet Eternity," *Phrack Magazine*, vol. 7, no. 51, 1997, http://www.cypherspace.org/eternity/phrack.html, retrieved Nov 2012.

[3] R. Barnes, A. Cooper, and O. Kolkman, "Technical Considerations for Internet Service Filtering," IETF-Draft, http://tools.ietf.org/html/draft-iab-filtering-considerations-01, 2012, work in progress.

[4] O. Berthold, H. Federrath, and S. Köpsell, "Web MIXes: A System for Anonymous and Unobservable Internet Access," in *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, H. Federrath, Ed. Springer-Verlag, LNCS 2009, July 2000, pp. 115–129.

[5] S. Burnett, N. Feamster, and S. Vempala, "Chipping Away at Censorship Firewalls with User-Generated Content," in *Proc. 19th USENIX Security Symposium, Washington, DC*, 2010.

[6] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, 1981.

[7] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, "Freenet: A Distributed Anonymous Information Storage and Retrieval System," in *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, July 2000, pp. 46–66.

[8] R. Clayton, S. J. Murdoch, and R. N. M. Watson, "Ignoring the Great Firewall of China," in *Proceedings of the Sixth Workshop on Privacy Enhancing Technologies (PET 2006)*, G. Danezis and P. Golle, Eds. Cambridge, UK: Springer, June 2006, pp. 20–35.

[9] G. Danezis and R. Anderson, "The Economics of Censorship Resistance," in *Proceedings of Workshop on Economics and Information Security (WEIS04)*, May 2004.

[10] R. Dingledine and N. Mathewson, "Design of a blocking-resistant anonymity system," The Tor Project, Tech. Rep. 2006-1, November 2006.

[11] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[12] K. P. Dyer, S. E. Coull, T. Ristenpart, and T. Shrimpton, "Peek-a-Boo, I Still See You: Why Efficient Traffic Analysis Countermeasures Fail," in *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, May 2012.

[13] N. Feamster, M. Balazinska, G. Harfst, H. Balakrishnan, and D. Karger, "Infranet: Circumventing Web Censorship and Surveillance," in *Proceedings of the 11th USENIX Security Symposium*, August 2002.

[14] N. Feamster, M. Balazinska, W. Wang, H. Balakrishnan, and D. Karger, "Thwarting Web Censorship with Untrusted Messenger Delivery," in *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, R. Dingledine, Ed. Springer-Verlag, LNCS 2760, March 2003, pp. 125–140.

[15] D. Fifield, N. Hardison, J. Ellithorpe, E. Stark, R. Dingledine, P. Porras, and D. Boneh, "Evading Censorship with Browser-Based Proxies," in *Proceedings of the 12th Privacy Enhancing Technologies Symposium (PETS 2012)*. Springer, July 2012.

[16] K. Fisher, "The death of SuprNova.org," http://arstechnica.com/staff/2005/12/2153/, December 2005, retreived Nov 2012.

[17] I. Goldberg and D. Wagner, "TAZ Servers and the Rewebber Network: Enabling Anonymous Publishing on the World Wide Web," *First Monday*, vol. 3, no. 4, August 1998.

[18] A. Houmansadr, T. Riedl, N. Borisov, and A. Singer, "IP over Voice-over-IP for censorship circumvention," *arXiv preprint arXiv:1207.2683*, 2012.

[19] A. Houmansadr, G. T. K. Nguyen, M. Caesar, and N. Borisov, "Cirripede: Circumvention Infrastructure using Router Redirection with Plausible Deniability," in *Proceedings of the 18th ACM conference on Computer and Communications Security (CCS 2011)*, October 2011.

[20] S. Hsu, "TriangleBoy," Whitepaper,

http://www.webrant.com/safeweb_site/html/www/tboy_whitepaper.html, 2000, retrieved Nov. 2012.

[21] ICANN Security and Stability Advisory Committee, “Impacts of Content Blocking via the Domain Name System,” http://www.icann.org/en/groups/ssac/documents/sac-056-en.pdf, October 2012, iCANN SSAC security advisory.

[22] L. Invernizzi, C. Kruegel, and G. Vigna, “Message in a bottle: Sailing past censorship,” in *Privacy Enhancing Technologies Symposium (PETS)*, 2012.

[23] Iran Daily Brief, “Arrested after interview on BBC and VOA on expected execution of brother,” http://iranbriefing.net/?p=15414, October 2012, retrieved Nov 2012.

[24] N. Jones, M. Arye, J. Cesareo, and M. J. Freedman, “Hiding Amongst the Clouds: A Proposal for Cloud-based Onion Routing,” in *Proceedings of the USENIX Workshop on Free and Open Communications on the Internet (FOCI 2011)*, August 2011.

[25] J. Karlin, D. Ellard, A. W. Jackson, C. E. Jones, G. Lauer, D. P. Mankins, and W. T. Strayer, “Decoy Routing: Toward Unblockable Internet Communication,” in *Proceedings of the USENIX Workshop on Free and Open Communications on the Internet (FOCI 2011)*, August 2011.

[26] S. Köpsell and U. Hillig, “How to Achieve Blocking Resistance for Existing Systems Enabling Anonymous Web Surfing,” in *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2004)*, Washington, DC, USA, October 2004.

[27] C. S. Leberknight, M. Chiang, H. V. Poor, and F. Wong, “A Taxonomy of Internet Censorship and Anti-Censorship,” http://www.princeton.edu/~chiangm/anticensorship.pdf, December 2010, draft.

[28] K. Levchenko and D. McCoy, “Proximax: Fighting Censorship With an Adaptive System for Distribution of Open Proxies,” in *Proceedings of Financial Cryptography and Data Security (FC’11)*, February 2011.

[29] P. Lincoln, I. Mason, P. Porras, V. Yegneswaran, Z. Weinberg, J. Massar, W. A. Simpson, P. Vixie, and D. Boneh, “Bootstrapping Communications into an Anti-Censorship System,” in *Proceedings of the USENIX Workshop on Free and Open Communications on the Internet (FOCI 2012)*, August 2012.

[30] J. McDonald, “China’s alternative to Twitter closes accounts over ’rumors’,” Associated Press,, retrieved Nov 2012.

[31] D. McPherson, “When Hijakcing the Internet,” http://ddos.arbornetworks.com/2008/11/when-hijacking-the-internet/, November 2008, retrieved Nov 2012.

[32] H. M. Moghaddam, B. Li, M. Derakhshani, and I. Goldberg, “SkypeMorph: Protocol Obfuscation for Tor Bridges,” in *Proceedings of the 19th ACM conference on Computer and Communications Security (CCS 2012)*, October 2012.

[33] G. Perng, M. K. Reiter, and C. Wang, “Censorship Resistance Revisited,” in *Proceedings of Information Hiding Workshop (IH 2005)*, June 2005, pp. 62–76.

[34] M. Schuchard, J. Geddes, C. Thompson, and N. Hopper, “Routing Around Decoys,” in *Proceedings of the 19th ACM conference on Computer and Communications Security (CCS 2012)*, October 2012.

[35] W. Scott, R. Cheng, J. Li, A. Krishnamurthy, and T. Anderson, “Blocking-Resistant Network Services using Unblock,” http://unblock.cs.washington.edu/unblock.pdf, October 2012, retrieved Oct. 2012.

[36] A. Serjantov, “Anonymizing censorship resistant systems,” in *Proceedings of the 1st International Peer To Peer Systems Workshop (IPTPS 2002)*, March 2002.

[37] Skype, “Skype,” http://www.skype.com/intl/en-us/home, retrieved Nov 2012.

[38] The New York Times, “Anonymous (Internet Group),” http://topics.nytimes.com/top/reference/timestopics/organizations/a/anonymous_internet_group/index.html, retrieved Nov 2012.

[39] ——, “Ecuador Grants Asylum to Assange, Defying Britain,” http://www.nytimes.com/2012/08/17/world/americas/ecuador-to-let-assange-stay-in-its-embassy.html?pagewanted=all&_r=0, August 2012, retrieved Nov 2012.

[40] ——, “Julian Assange,” http://topics.nytimes.com/top/reference/timestopics/people/a/julian_p_assange/index.html?inline=nyt-per, August 2012, retrieved Nov 2012.

[41] The Wall Street Journal, “Google Outage Shows Business Risks in China,” http://online.wsj.com/article/SB10001424127887324073504578112733488674060.html, November 2012, retrieved Nov 2012.

[42] Tom Online Inc., “TOM Skype,” http://skype.tom.com/, retrieved Nov 2012.

[43] Tor Project Inc., “Experimental Defense for Website Traffic Fingerprinting,” Tor Blog, https://blog.torproject.org/blog/experimental-defense-website-traffic-fingerprinting, September 2011, retrieved Nov 2012.

[44] ——, “Iran blocks Tor; Tor releases same-day fix,” https://blog.torproject.org/blog/iran-blocks-tor-tor-releases-same-day-fix, September 2011, retrieved Nov 2012.

[45] ——, “Obfsproxy: the next step in the censorship arms race,” Tor Blog, https://blog.torproject.org/blog/obfsproxy-next-step-censorship-arms-race, February 2012, retrieved Nov 2012.

[46] ——, “Top changes in Tor since the 2004 design paper (Part 2),” Tor Blog, https://blog.torproject.org/blog/top-changes-tor-2004-design-paper-part-2, October 2012, retrieved Nov 2012.

[47] M. Waldman and D. Mazières, “Tangler: A Censorship-Resistant Publishing System based on Document Entanglements,” in *Proceedings of the 8th ACM Conference on Computer and Communications Security (CCS 2001)*, November 2001, pp. 126–135.

[48] M. Waldman, A. Rubin, and L. Cranor, “Publius: A robust, tamper-evident, censorship-resistant and source-anonymous web publishing system,” in *Proceedings of the 9th USENIX Security Symposium*, August 2000, pp. 59–72.

[49] Q. Wang, X. Gong, G. T. K. Nguyen, A. Houmansadr, and N. Borisov, “CensorSpoofer: Asymmetric Communication using IP Spoofing for Censorship-Resistant Web Browsing,”

in *Proceedings of the 19th ACM conference on Computer and Communications Security (CCS 2012)*, October 2012.

[50] Z. Weinberg, J. Wang, V. Yegneswaran, L. Briesemeister, S. Cheung, F. Wang, and D. Boneh, "StegoTorus: A Camouflage Proxy for the Tor Anonymity System," in *Proceedings of the 19th ACM conference on Computer and Communications Security (CCS 2012)*, October 2012.

[51] wikileaks.org, "What is Wikileaks?" http://wikileaks.org/About.html, retrieved Nov 2012.

[52] B. Wiley, "Blocking-Resistant Protocol Classification Using Bayesian Model Selection," 2011.

[53] ——, "Dust: A Blocking-Resistant Internet Transport Protocol," *Technical report. http://blanu.net/Dust.pdf*, 2011.

[54] S. Wolchok, R. Yao, and J. A. Halderman, "Analysis of the Green Dam censorware system," *Computer Science and Engineering Division, University of Michigan*, vol. 18, 2009.

[55] E. Wustrow, S. Wolchok, I. Goldberg, and J. A. Halderman, "Telex: Anticensorship in the Network Infrastructure," in *Proceedings of the 20th USENIX Security Symposium*, August 2011.