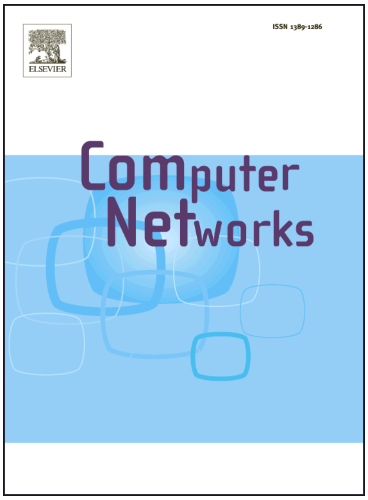

Internet Censorship Detection: a Survey

Giuseppe Aceto, Antonio Pescapé

PII: S1389-1286(15)00094-8
DOI: http://dx.doi.org/10.1016/j.comnet.2015.03.008
Reference: COMPNW 5540

To appear in: *Computer Networks*

Received Date: 22 September 2014
Revised Date: 27 February 2015
Accepted Date: 9 March 2015

Please cite this article as: G. Aceto, A. Pescapé, Internet Censorship Detection: a Survey, *Computer Networks* (2015), doi: http://dx.doi.org/10.1016/j.comnet.2015.03.008

Internet Censorship Detection: a Survey

Giuseppe Aceto and Antonio Pescapé
University of Napoli Federico II (Italy)
{giuseppe.aceto, pescape}@unina.it

***Abstract*—Internet Censorship is a phenomenon that crosses several study fields, from computer networking and computer security to social sciences; together with censorship detection and censorship circumvention it has impact on Internet infrastructure, protocols and user behaviors. Detection of Internet Censorship is the basis for the study of this phenomenon, and recently it has received focus from a technical point of view. Due to the heterogeneity of study fields and approaches, the scientific corpus on these topics is still in need of an overall analysis, based on coherent framework and lexicon to describe the experimented approaches and findings.**

In this paper we present a survey on Internet Censorship detection. We propose a reference for censoring techniques and a characterization of censoring systems, with definitions of related concepts. Based on the censoring techniques investigated in literature, we propose an analysis and discussion of censorship detection techniques and architectures and we present a chronological synopsis of the literature adopting or introducing them. Then we collect, study, and discuss available tools and platforms for censorship detection, and propose a characterization scheme to analyze and compare them. Finally, we compare and discuss detection architectures, tools and platforms, and we use the results to infer current challenges and for proposing new directions in the field of censorship detection.

***Index Terms*—Internet Censorship, Network Monitoring, Communications Surveillance, Privacy, Network Security.**

1 Introduction and Motivations

Internet Censorship is a complex phenomenon that is deeply discussed and analyzed in the field of social sciences, and in recent years has attracted attention also from other study fields such as computer security and computer networking due to the widespread adoption of ICT for information control, previously focused on analog mass media. Putting aside the social and political aspects evidently related to censorship, we focus on the technical aspects only: regardless of the aims, scope or legitimacy of it, we consider “Internet Censorship” as the intentional impairing or blocking of access to online resources and services. The design principles of the Internet as an open and distributed system contrast with the controls required by censorship. Therefore the technical means adopted to this end almost invariably imply the interference with—or disruption of—standard network protocols and expected behavior of network applications. This has practical consequences for all the stakeholders of the Internet: the end users, which are subject to restrictions and impairments with varying degrees of transparency; the ISPs, that face the complicated trade-off among complying with the law, building and managing the censorship infrastructure, and providing the best service to their customers; the transit operators, that potentially experience unexpected traffic patterns; finally the online service providers, that may have to deploy and operate censorship systems as demanded by the law of their own country, and whose global user base (up to whole countries at a time) can be subjected to impairments or complete blocking. Moreover, due to the complexity and the non-standard nature of censorship techniques, unforeseen side effects can strike third parties (as actually already happened [110]). Summarizing, even if adopted for legitimate and embraceable aims censorship requires mangling of several components of the Internet and this has an impact on all its stakeholders. Several systems for *circumvention* and *detection* of Internet Censorship have been developed over the years; these too are of interest for the different Internet stakeholders, according to their roles. In fact surveillance needs to recognize the related traffic, and prevent both false negatives (when *circumvention* is effective) and false positives (when *censorship detection* triggers the alarms); users and online service providers may be interested in *circumvention* techniques to prevent side effects or unlawful interference (besides illicitly eluding the restrictions). In addition to the aforementioned reasons, censorship *detection* in its turn is of central importance for different actors. For academy and industry researchers, the study and employment of censorship *detection* is functional to understanding if, to what extent, and with which method censorship is enforced. Significant aspects of censorship, such as its enforceability, effectiveness, and transparency, as well as the possible unwanted side effects, strongly depend on the technical details of the adopted censorship technique and thus evolve with the technology and real usage of it. For the creators of circumvention systems, the mechanics of censorship revealed by *detection* are at the basis of the design and development of their tools. Finally, for the operators and users that perform network diagnostics, the *detection* of censorship can provide the explanation for apparent outages and malfunctioning, discharging the inculpable application, network administrator, ISP, or online service provider. In brief, for many different actors censorship *detection* is either very valuable or strictly necessary.

Despite this, to the best of our knowledge, no peer-reviewed survey has investigated such topics, moreover no survey is available that specifically addresses Internet Censorship *detection*. Previous works have surveyed and analyzed Internet Censorship and circumvention techniques and tools (Leberknight et al. [99], Elahi and Goldberg [50]). An analysis of the different phases of the application of Internet Censorship and the citizens’ perception and reaction to it is presented by Bambauer [16]. The studies on Internet Censorship and papers proposing detection and circumvention techniques often report a technical analysis of selected censorship techniques and related works; these however are not meant to be surveys, and