TOWARDS AN INFORMATION CONTROL POLICY MODEL
Achieving More Transparency in Internet Filtering Approaches

Andreas Kasten
Institute for IS Research, University of Koblenz-Landau, Universitätsstraße 1, 56070, Koblenz, Germany

Keywords: Information flow control, Internet communication, Internet filtering, Internet control, Policy language model.

Abstract: Internet filtering is the manipulation of Internet communication in order to prevent the access and exchange of unwanted data. According to the reports of the OpenNet Initiative, Internet filtering emerges all over the world. Although many filtering techniques are legitimated by a similar legal basis, most of them are implemented differently. This paper explains the need for a policy language model that is able to describe Internet filtering techniques on different levels of abstraction including their legal basis and their technical implementation aspects. The paper further explains the requirements for such a model and outlines a first concept.

1 INTRODUCTION

The *OpenNet Initiative* (ONI) is a collaboration of three institutions that gather information about Internet filtering and surveillance in several countries. The initiative analyzes the technical implementations of different filtering and surveillance techniques and also investigates their impact on the civil society. The ONI regularly publishes its results such as in (Deibert et al., 2008) and (Deibert et al., 2010). According to these reports, state-mandated Internet filtering emerges all over the world. However, not all countries provide an adequate explanation about the details of or the reason for their filtering (Faris and Villeneuve, 2008). For example, although both China and Saudi Arabia practice Internet filtering, they essentially differ in respect of their level of transparency (Zittrain and Palfrey, 2008a). Furthermore, the technical implementations of filtering often differ between countries, although they may sometimes be legitimated through similar legal bases. Thus, (Zittrain and Palfrey, 2008b) ask for more transparency in Internet filtering approaches. This request mainly addresses international corporations in the information and communication technology sector, since these corporations often carry out state-mandated Internet filtering.

The *Global Network Initiative* (GNI) consists of such corporations as well as universities and other non-governmental organizations. The GNI aims at protecting the freedom of expression and the privacy of Internet users in every country, independently of how elaborately the country is controlling the Internet. The GNI created and published its *Principles on Freedom of Expression and Privacy* (GNI, 2011) that define how GNI members are supposed to act when it comes to Internet filtering and surveillance. GNI members shall inform Internet users about the country's laws and how they carry out these laws. However, the principles only cover the organizational level and do not give any technical implementation guidelines.

This paper outlines the basic aspects of a comprehensive policy language model that covers the legal and the organizational level of Internet communication manipulation and the details of its technical implementation. This policy model aims at providing a framework for describing several types of Internet manipulations. Such a framework may be used as a communication medium between different countries and corporations that participate in Internet filtering. The policy model shall therefore be able to prevent different interpretations and enforcements of the same legal basis. It can be considered as an extension to the principles of the GNI by following the results of the ONI and the requests made by (Zittrain and Palfrey, 2008b). Furthermore, by describing both organizational and technical aspects of Internet manipulations, the policy model shall make these manipulations more transparent to Internet users.

The rest of this paper is structured as follows: Section 2 briefly explains some basics of Internet

Kasten A..
TOWARDS AN INFORMATION CONTROL POLICY MODEL - Achieving More Transparency in Internet Filtering Approaches.
DOI: 10.5220/0003607804090413
In *Proceedings of the International Conference on Security and Cryptography* (SECRYPT-2011), pages 409-413
ISBN: 978-989-8425-71-3

filtering. Section 3 outlines similar approaches of policy based communication filtering. Section 4 explains some basic requirements for the proposed policy language model. Section 5 outlines an abstract concept for such a model. Finally, section 6 summarizes the main ideas with a short conclusion.

2 INTERNET FILTERING BASICS

Every Internet connection is established between a source and a destination system. These two endpoints don't normally communicate directly with each other but rather via several other components such as routers, DNS servers, or even proxy servers. All these components including the endpoints systems can generally be used for manipulating the communication channel.

Each component is operated by a specific party, which in turn is able to configure the component in respect of its manipulating functions. Following (Zittrain, 2003) and (Deibert and Villeneuve, 2004), a component can be associated with one of five different parties: one owner of each endpoint system, one Internet service provider (ISP) for each owner, and one abstract party that operates the rest of the Internet such as backbones. The endpoint systems may be owned by a private person, a cyber cafe, or a public institution such as a school or a library. ISPs may not only provide access to the Internet but also a web hosting service. Backbones may be operated by a country's government.

All of these five parties can manipulate an Internet connection and even prevent its establishment. In other words, all of them are able to perform a specific kind of Internet filtering. The possible filtering techniques depend on the general functions of the filtering component. (Dornseif, 2004), (Clayton, 2006) and (Murdoch and Anderson, 2008) identify three main categories of Internet filtering techniques based on the type of information that they require.

IP packet dropping mechanisms are based on the information of the Internet protocol's header data. All packets to or from a specific IP address or address range are dropped instead of forwarding them. The techniques of this category are basically supported by every router. Filtering techniques based on the *DNS protocol* affect the mapping from a domain name to a corresponding IP address. Instead of returning the correct IP address, a DNS server application returns either an incorrect IP address or none at all. *Content analysis* techniques require a further evaluation of the actual content being exchanged and highly depend on the content's type. They may be based on URL comparison, keyword matching, or even image recognition. As these techniques often require the data of the application layer, they are often implemented using a proxy server. *Hybrid filtering techniques* combine different approaches. (Clayton, 2006) studied a filtering system that operates in two steps. First, a router looks for suspicious IP packets. These packets are then forwarded to a proxy server, which deeper analyzes their content and makes a final filtering decision.

3 POLICY BASED CONTROL

This section outlines different approaches for policy based control mechanisms. Although the mechanisms have slightly different foci, they are all able to describe specific aspects of Internet filtering.

The *Platform for Internet Content Selection* (PICS) (Resnick and Miller, 1996) defines a labelling framework for annotating web resources. The framework provides a format for creating labels and linking them to web resources, and specifies how the labels can be accessed by an Internet user. PICS does not specify any labelling vocabulary and leaves the creation of such to its users. Furthermore, it does neither define by whom the labels shall be created nor how they shall be evaluated. PICS is flexible enough to be implemented in client-based filtering software or even within proxy servers.

The *Platform for Privacy Preferences Project* (P3P) (Cranor, 2003) specifies a format for describing privacy policies of web sites that can be automatically interpreted by a web browser. This shall help end users to easier understand a web site's privacy policy. In order to use a P3P policy, a user has to specify her privacy preferences. When she visits a web site, her client application downloads the P3P policy of that web site, checks it against her privacy preferences and notifies her about the result. However, P3P policies are not designed for direct technical enforcement (Anderson, 2005).

The *Enterprise Privacy Authorization Language* (EPAL) (Ashley et al., 2003) defines a language model for specifying and enforcing privacy policies of corporations. Such policies describe how a party may use what personal data for what purpose. Contrary to P3P, EPAL does not want to achieve more transparency for end users. Instead, it focuses on enforcing existing privacy policies by translating them into a technical description. This description can also be used as a common interchange format between different corporations that process the same personal data. Since EPAL does not define any specific vocabulary, a corporation must create its own.

The *Extensible Access Control Markup Language* (XACML) (Moses, 2005) defines a framework for policy based access control. This framework includes a rule-based policy language, a format for authorization messages, and an architecture for processing the policies. XACML is rather similar to EPAL, but has a much broader focus and greater expressiveness (Anderson, 2005). XACML is not specifically designed for privacy policies, but defines an open framework that can be used for different access control implementations. XACML's policy language can even be extended with additional language elements.

The *Open Digital Rights Language* (ODRL) (Iannella, 2002) is a rights expression language for describing rights over physical or digital goods. It can be used as a general interchange format between different DRM systems. ODRL defines a basic language model and a data dictionary. It allows for describing which parties can perform which actions on which assets. An application using ODRL may use ODRL's default data dictionary or create its own. The creation of a specific ODRL profile for P3P is also possible. ODRL only focuses on a general language model and a vocabulary. It neither defines an implementation nor an interpretation guideline for specific ODRL licenses.

4 REQUIREMENTS FOR A POLICY LANGUAGE MODEL

A policy model for describing Internet filtering must be able to describe its legal, organizational, and technical level. The first two levels allow for a greater transparency and background knowledge of the filtering methods whereas the latter level provides the parameters for its technical enforcement.

Consider an example: Italian ISPs are required by law to block access to unlicensed gambling web sites (Deibert et al., 2010). A policy model should be able to describe the law itself, its integration into an ISP's code of conduct, and how the ISP technically implements the filtering. None of the policy models of section 3 is able to appropriately describe all three aspects. EPAL and XACML focus on a more or less predefined environment of corporations that act according to their agendas, but are not able to describe the legal bases of Internet filtering in detail. P3P and ODRL do not consider the enforceability of their policy models. PICS can generally be used for filtering, but lacks a formal representation of its labels.

The basic requirements for a policy model that covers all three aspects are transparency, enforceability, expressiveness, expandability, and the ability to allow different views on the filtering process. As explained in section 1, transparency is one of the main requirements for the proposed policy model. It covers the *legal basis* for the filtering, the *type of the filtered information*, and the *filtering parties*. The legal basis authorizes the filtering mechanism. The description of this basis allows for a better understanding why the filtering is carried out in the first place. Since a code of law may be too abstract concerning the actual type of the filtered content, a specific policy should describe this content in further detail. Naming the filtering parties within such a policy provides contact information for asking further questions about the filtering process.

In order to be enforceable via technical components, a specific policy must describe the used filtering function and its parameters. For example, if IP packet dropping shall be used, the policy must describe both this method and its input parameters. In this case, the input parameters are IP addresses or address ranges. In order to prohibit misinterpretations and different implementations of a specific policy, the policy model must allow for detailed and unambiguous descriptions. This requirement may be achieved via the use of a formal language.

The policy model shall provide a general framework for describing and communicating different filtering techniques. It must therefore be able to express already existing filtering implementations. As shown in section 2, Internet filtering can be carried out by different technical components. The policy model must therefore be able to describe such components as well as different filtering implementations. For example, the model must be able to both describe IP packet dropping as well as filtering methods based on the DNS protocol.

The proposed model must be interoperable with already existing policy models such as those of section 3. Furthermore, it should be expandable in order to also cover future policy models. It should therefore provide an open interface that allows for adding further language elements. In order to ease expandability, the policy model's internal structure should be modular. A modular policy model allows for creating new sub-models that expand the expressiveness of the base model (Scherp et al., 2011).

Transparency is a core requirement for the policy model and shall be accomplished on several levels of abstraction. However, the more details about a filtering mechanism are known, the easier it can be circumvented. In order to achieve transparency and prevent circumvention at the same time, the policy model should provide different views for a specific policy. For example, the filtering party must access all details of the filtering mechanism including its general functioning and its required parameters. On the other hand, regular Internet users must not see all these details, but they must have access to the legal basis that authorizes the filtering.

5 AN ABSTRACT CONCEPT FOR A POLICY MODEL

This section outlines a possible structure of a policy model that satisfies the requirements of section 4. The proposed structure consists of three layers of different expressiveness and abstraction. The first layer is the most general one and only provides non-technical information. The other two layers expand their preceding layers by adding further details.

The first layer contains information about the legal bases for the Internet filtering. It refers to the specific statutes that legitimate the filtering, briefly describes their contents, and states the topics of the data to be filtered. Since this layer only contains public information, there are no constraints on its accessibility. Its contents may be directly provided by the filtering country's government.

The second layer outlines the code of conduct of the filtering party. This code of conduct extends the legal basis with party-specific regulations. In most cases, this party is a corporation acting on behalf of the state it operates in such as the members of the GNI. The second layer also contains only abstract information about the filtering and can therefore be accessed by any party. Its contents are directly provided by the filtering party. The first and the second layer satisfy the transparency requirement.

The third layer extends the abstract regulations of the first two layers with technical implementation details. These cover the filtering components, the specific methods, and their required input parameters. An example dataset of this layer contains the IP addresses that are used for IP packet dropping. Since this layer contains sensitive information concerning the circumvention of the filtering, its access must be restricted to the filtering parties. Furthermore, this layer is necessary as a communication tool between different filtering parties that act on behalf of the same statues. It conforms to the enforceability requirement.

6 CONCLUSIONS

This paper explained the need for a policy language model for Internet communication filtering as stated by the reports of the ONI and the principles of the GNI. The paper also outlined the basic requirements for such a policy model and outlined an abstract concept which fulfils these requirements. The specific details of the proposed model and its formalization must still be developed. In order to achieve a broader reusability of the model, it is intended to design the model as a web ontology. The expressiveness of the policy's layered structure could also be extended. A fourth layer describing the user's view could be added. This layer could describe the contents the user wants to access, the legal bases she is bound to, and the technical components she uses for her Internet communications.

REFERENCES

Anderson, A., 2005. Comparison of Two Privacy Policy Languages: EPAL and XACML. *Sun Microsystems Laboratories.*

Ashley, P., Hada, S., Karjoth, G., Powers, C., Schunter, M., 2003. Enterprise Privacy Authorization Language (EPAL 1.2). *W3C Member Submission*.

Clayton, R., 2006. Failures in a Hybrid Content Blocking System. In *Proceedings of the PET Workshop*.

Cranor, L. F., 2003. P3P: Making Privacy Policies More Useful. In *IEEE Security and Privacy*, vol. 1(6).

Deibert, J., Villeneuve, N., 2004. Firewalls and Power: An Overview of Global State Censorship of the Internet. In Human Rights in the Digital Age. *GlassHouse Press*.

Deibert, J., Palfrey, J., Rohozinski, R., Zittrain, J., 2008. Access Denied: The Practice and Policy of Global Internet Filtering. *The MIT Press*.

Deibert, J., Palfrey, J., Rohozinski, R., Zittrain, J., 2010. Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace. *The MIT Press*. Springer.

Dornseif, M., 2004. Government mandated blocking of foreign Web content. In *Security, E-Learning, E-Services: Proceedings of the 17. DFN-Arbeitstagung über Kommunikationsnetze*.

Faris, R., Villeneuve, N., 2008. Measuring Global Internet Filtering, In *Deibert*, 2008, pp. 5-27.

Iannella, R., 2002. Open Digital Rights Language (ODRL) Version 1.1. W3C Note.

Global Network Initiative (GNI), 2011. Principles on Freedom of Expression and Privacy.

Moses, T., 2005. eXtensible Access Control Markup Language (XACML) Version 2.0. *OASIS Standard*.

Murdoch, S. J., Anderson, R., 2008. Tools and Technology of Internet Filtering. In *Deibert,* 2008, pp. 57-72.

Resnick, P., Miller, J., 1996. PICS: Internet Access Controls Without Censorship. In *Communications of the ACM*.

Scherp, A., Saathoff, C., Franz, T., Staab, S., 2011: Designing Core Ontologies. University of Koblenz-Landau.

Zittrain, J., 2003. Internet Points of Control. Berkman Center for Internet & Society.

Zittrain, J., Palfrey, J., 2008a. Internet Filtering: The Politics and Mechanisms of Control. In (Deibert, 2008), pp. 29-56.

Zittrain, J., Palfrey, J., 2008b. Reluctant Gatekeepers: Corporate Ethics on a Filtered Internet. In (Deibert, 2008), pp. 103-122.