Privacy-Aware Secure Monitoring

Sathya Rao[1], Giuseppe Bianchi[2]
Ivan Gojmerac[3], Georgios Lioudakis[4]
[1] Telscom AG, Switzerland, Rao@Telscom.ch
[2] University of Rome, Italy, giuseppe.bianchi@uniroma2.it
[3] Telecommunications Research Center Vienna (ftw.), Austria, gojmerac@ftw.at
[4] National Technical University of Athens, Greece, gelioud@icbnet.ntua.gr

Abstract: Traffic monitoring is necessary for the operation, maintenance and control of communication networks. Traffic monitoring has also important implications on the user privacy. This paper discusses a novel approach to privacy-preserving traffic monitoring and the related research challenges. This study is based on the analysis of the technical implications and regulatory provisions in the areas of data protection and security. We propose a two-tier privacy-preserving monitoring architecture characterized by i) the adaptation and operation of monitoring applications on protected data, and ii) a back-end middleware system devised to control and orchestrate the access to and processing of the collected data. The feasibility of the proposed architecture faces many challenges. However, by carefully designing data protection mechanisms, by adapting monitoring applications, and by deploying a semantic-rich privacy-aware access control framework, it is possible to concurrently meet strong privacy requirements, achieve efficient solutions for traffic monitoring, and be compliant with the recent regulatory provisions such as data retention.

Keywords: privacy, network monitoring, privacy legislation, access control, semantic model.

1. Introduction

Privacy is of great concern to users of the Internet, and is a critical part of a user experience. The PRISM project investigates the possibility to preserve the customers' privacy, by avoiding disclosure of raw captured data even inside the controller domain itself, while preserving the possibility of running monitoring applications, including the possibility to detect and react to attacks and trace back abuses (thus improving public security). The PRISM technology aims at being fully legally compliant with data privacy protection regulation on one side, and to the security legislation on the other side.

The Goal of the PRISM project is to devise network monitoring technologies and architectures, which guarantee enforcement of data protection legislation. This will be accomplished through the specification, design, implementation and validation of a two-tiered network monitoring system. The design of a comprehensive monitoring system cannot be considered as a purely technical activity, as it affects the society as a whole (in terms of both its implications on privacy and in terms on its consequences on public security). Moreover, European policies and regulatory requirements do affect to a significant extent the technical requirements of a network monitoring system.

2. PRISM System Architecture

The key engineering guidelines for privacy-preservation are:

- Protect the data as soon as they are captured, i.e. on the on-line monitoring probes
- Adapt monitoring applications to operate on protected data
- Decouple the entity in charge of enforcing data protection (e.g. legal authority) from that one in charge of running, e.g. monitoring applications (i.e., the network operator)
- Provide a comprehensive framework for the control of the access and processing of the stored data traces.

These ideas are reflected in the multi-component, two-tiered system architecture sketched in Figure 1. Unlike traditional architectures, that are typically monolithic from the functional point of view, the envisioned system is comprised of three separate sub-systems that are also administratively independent: the Front-End tier, the Back-End tier and the Privacy-Preserving Controller (PPC).

PRISM system architecture has four functional blocks:

PRISM Front-end – This component is meant to be a "black-box" traffic probe, "cryptographically controlled" by an entity, in the figure referred to as third-party privacy-preserving controller. The PRISM front-end is devised to capture data on the network link(s), protect them according to suitably designed data protection mechanisms whose secrets are provided by the Privacy-Preserving Controller, and deliver them to the back-end system through standard-based data export protocols, IPFIX being the technology of choice.

Privacy-Preserving Controller – This entity accomplishes the task of providing and maintaining the crypto secrets, which are used by the data protection mechanisms enforced on the front-end.

PRISM back-end – This part of the system is in charge of collecting, storing and processing the front-end protected data traces. Monitoring applications running on the back-end will operate on encrypted traces, and when strictly necessary and/or mandated by regulatory provisions. It will interoperate with the privacy preserving controller to selectively revert the data protection mechanisms set forth at the front-end.

Public Domain – Finally, collected data traces and/or derived statistics will be further sanitised through robust anonymization mechanisms. These will allow disclosure of data traces and/or related derived information to the public community, to meta-repositories, and to externally operated monitoring applications.

3. Adaptation of Monitoring Applications

The challenge of the architecture is the possibility to operate monitoring applications in presence of data protection mechanisms, with zero or minimal loss of their effectiveness. These applications rely on protocol information that is available in the packet headers and in the initial part of the payload and they do not require complete access to the packet payload and the privacy problem reduces to the protection of the identity associations (anonymization). The security-oriented applications and traffic classification will use deep packet inspection and relies on performing fine-grain (bit-level) analysis of the whole

payload. These are mainly signature-based Intrusion Detection Systems (IDS). Resolving the privacy problem is extremely tough in this case as full payload access is an intrinsic requirement.

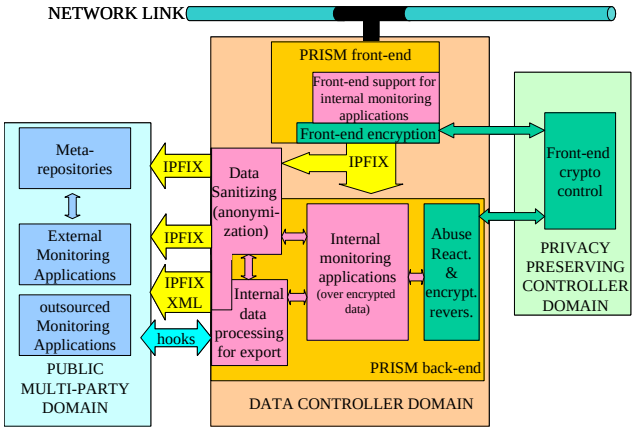

Figure 1. PRISM system architecture

4. A Semantic Model for the Back-End middleware

Privacy-aware access control constitutes a very important feature of the back-end tier. It provides a generalization of the notion of "access" to stored data, where the data access requests are evaluated against a dynamic "context" state characterized by a multiplicity of parameters, including i) the type and identification of the user or requesting entity, ii) the type of the data to be accessed, iii) the processing intended, iv) a stated purpose, v) the environment on which the data will be used, vi) the applicable regulatory provisions for the requested data and requesting entity types, vii) other provisions such as a "consulting" verdict of a Bayesian filter about the access request, etc.

The back-end system is empowered with a policy-based decision engine that takes into account the privacy requirements. This policy engine reacts to access requests based on a dynamic "privacy context" that incorporates all parameters into a single evaluation block. Matching the privacy context against a model of defined privacy-related access policies enables the back-end system to take a decision and apply the appropriate protective measure before releasing the data. The model of defined policies is implemented by means of an ontology. It provides a formal and machine-readable representation of policy rules that originate from a semantic description of the monitoring applications operation and that explicitly includes privacy legislation provisions. The ontology constitutes the core of data protection at the back-end tier. It provides the configuration of the decision engine and defines how the back-end tier will treat the data before exposing them to the requesting entity.

Based on the decision engine outcome, the back-end tier mediates access to the stored data by the execution of a processing function or a combination of multiple modules. An immediate consequence of this approach is the ability to support third-party "outsourced" monitoring applications. In principle, monitoring applications developed by third-parties might require access to the internal captured data in order to perform advanced operations, which a robust anonymization mechanism devised for data public export would impede. By allowing the controlled execution of data processing procedures inside the controller domain, the back-end tier enables outsourced applications to operate on privacy-sensitive data. With this approach only the application output is accessed by the external party, not the input data. In this way it might be possible to define collaboration models for data processing between the network operators and external parties such as research groups in different universities.

5. Conclusions

This paper addresses the challenges and issues emerging in the design of privacy-aware network monitoring systems. The solution is seen in a two-tier architecture system. A front-end tier of data protection mechanisms is directly enforced at the traffic probe device, guaranteeing that the data delivered to the back-end storage will be already privacy-protected, while the back-end tier enforces privacy-aware access control to the collected data, orchestrating at the same time the operation of reversing the data protection mechanisms set forth by the front-end tier. The back-end system retains the ability to reverse the data protection measures by cooperating with an external element called "Privacy-Preserving Controller".

The approach proposed in this paper aims at satisfying three conflicting aspects: i) enable legal investigations by law enforcement authorities, ii) safeguard the right of the individual to communication privacy, and iii) allow the network operators to perform traffic monitoring as part of the operation process of their infrastructures.

Acknowledgements The authors acknowledge the PRISM consortium members for their contribution to develop this paper. The PRISM project is partly financed by the European Commission.