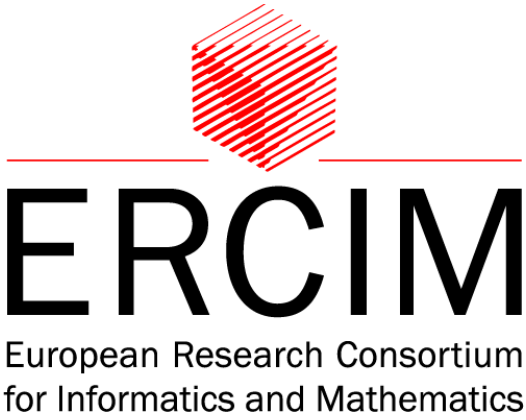

ERCIM White paper on Cyber-security and privacy research

Version v0.91, Oct. 2014

Contents

Introduction 3
Methodology 4
Structure 4
Application Domains 5
Open Web Platforms 5
Cloud 5
Social networks 6
Mobile 6
IoT 6
Smart Grid 7
Intelligent Transportation Systems 7
Research areas 9
Research Area: Systems security 9
Research Area: Security Engineering 12
Research Area: Security operation management 15
Research Area: Data Protection 18
Research Area: Security and Big Data 21
Research Area: Access Control 24
Research Area: Quantitative Aspects of Security 26
Research Area: Practical and Usable Security 29
Research Area: Cryptography 31
Research Area: Trust Management Systems 34
Research Area: Digital Freedom 37
Research Area: Network Security 39
Other aspects 43
Other aspects: Escaping the Legal Trap 43
Other aspects: “Crash Commission” for cyber? 46
Other aspects: Improving Awareness 48
Conclusion 49
Contributors 50

Introduction

This document describes the main research challenges on security and privacy identified by a group of experts of the European Research Consortium in Informatics and Mathematics (ERCIM).

In early 2014, ERCIM launched an initiative to identify the emerging of grand challenges in ICT and define the strategic research topics needed in order to achieve the highest impact results to meet those challenges. This kind of activity is in line with other similar and successful initiatives lead by ERCIM, also in cooperation with other institutions as EU.

In particular, the ERCIM Board of Directors (BoD) identified two main research fields in ICT, i.e. "big data analytics" and "security and privacy". For each of them, a group of expert has been identified and asked to produce a white paper identifying the research topics to be addressed.

This document summarizes the contributions by the "security and privacy" group of experts, mainly stemming from the ERCIM Security and Trust Management WG (ERCIM STM), representing an established think-thank in the area managed by ERCIM.

The initiative is led on behalf of this WG by Pierangela Samarati, Javier Lopez and Fabio Martinelli (current chair and previous ERCIM STM WG chairs, respectively). The ERCIM STM WG members where pleased to see that security and privacy are recognized as main research topics where ERCIM can express a critical mass.

Cyber-security and privacy are two major concerns of the scientific community as well as of today's globally connected society. The enormous advances in ICT allows us to enjoy ubiquitous and pervasive technologies that are at the very heart of almost every activity we perform. Together with these advantages come however new security and privacy risks and worries.

Our every-day life depends very much on ICT, from mobile phone usage (more than one billion of devices currently with the same operating systems), computers in offices and programmable machines in factories, and intelligent surveillance cameras contributing to our safety (and tampering our privacy).

This pervasiveness of ICT increases the potential attack surface exploitable by attackers, expanding the opportunities and potential damages of the exploits. The number and motivation of attackers are increasing, making cyber-security attacks a wealthy market.

Cyber-security is a long-standing research topic with many success stories in the literature and in the industry standards and products. Yet, cyber-attacks are increasing and their impact even more significant. This is definitely related to the expansion of ICT role and the fact that technologies evolve together their threats.

Cyber-attacks are continuously receiving attention in the media, thus raising the awareness and the concern on the society. The protection of personal information is also a main concern. Many consider this already a lost battle since most of our everyday life can be monitored by private companies (in addition to states). Lack of privacy should however not be a price to pay to enjoy technology and work is needed to enable users to enjoy technological advances while not giving up their privacy. Technological solutions should then be developed to empower users with full control over their own data as well as to provide technological support to legislations for the protection of data.

Confirming their importance, cyber-security and privacy are really active areas of research and experts in this field in growing demand by companies.

This document aims to help identifying research areas that could provide significant contributions to reduce the cyber-insecurity and where ERCIM Institutions have significant research and innovation capabilities able to drive cooperative research efforts in the field.

Methodology

In accordance to the initial appointment of the ERCIM BoD, the process to structure the white paper is based on expert group workshops and consensus building. Thus, we planned the focus expert group workshop on Sept. 9 2014, prior the ERCIM STM WG meeting (Sept. 10-11 2014) in Wroclaw, in cooperation with the 19th edition of the European Symposium on Research in Computer Security (ESORICS).

Before the workshop, we organized the collection of short position statements from members of the expert group, that were asked to briefly address the following issues:

- *which are the main threats to security and privacy in the next 2-5 years;*
- *which are the main research challenges and gaps (and why);*
- *which are the two top priorities to be addressed at European level.*

During the meeting, based on the positions statements received and the follow-up discussion, a set of relevant technological areas were identified, together with applications domains and other not technical aspects.

The initial findings of the expert group were presented during the ERCIM STM workshop and additional feedback on the research challenges to be addressed were received. Follow-up meetings were carried out via teleconference within the expert group members.

The present document has been then prepared by the expert group members who focus on their own area of expertise, trying to share the workload as well as the work ownership of the results. It is worth noticing that all the research areas identified represent areas of competence of two or more research organizations of the ERCIM consortium.

Structure

This document is structured as follows. Section 2 identifies the main application domains where cyber-security and privacy have a major role. Section 3 depicts the main research areas to be addressed, mainly from a technological perspective, giving the current state of the art and identifying several research problems to be solved. Section 4 emphasises on non-technological aspects that are to be taken into account for research on privacy and security, such as legal, regulatory and societal aspects. Section 5 gives concluding remarks.

October 2014

Javier Lopez

Fabio Martinelli

Pierangela Samarati

Application Domains

ICT are pervasive enablers for many other technologies, thus security of ICT is actually of paramount importance for the security of many application domains.

We highlight several application domains that would benefit further analysis and investigation. In particular:

1. Open Web platforms
2. Cloud
3. Social networks
4. Mobile devices
5. Internet of Things (IoT)
6. Smart Grid
7. Intelligent transport systems

More into the details.

Open Web Platforms

Despite the availability of a multitude of solutions and standards for securing web applications and services, the web still remains a vulnerable environment, subject to a multitude or threats and risks. Among them, user-to-service authentication, device-to-user authentication, web application vulnerabilities (e.g. XSS and SQL injection) stand out.

Research efforts in this domain should be directed towards developing authentication mechanisms beyond the traditional (and mostly used today password scheme); developing stronger, more efficient and easily deployable device-to-user authentication techniques; extending XACML policies to integrate credential support, context representation, and exception management; and towards practical ways of removing long standing vulnerabilities from existing popular and widely used services.

Cloud

Cloud computing constitutes a paradigm shift in the computing model, from an on-premise IT infrastructure to an off-premise IT service. As most of the security controls and mechanisms are closely associated with the infrastructure, the shift also results in security moving further away from both the user and the corporate IT manager. This, in turn, results in new threats and risks, different in nature and/or intensity than those encountered when infrastructure, platforms or software are not shared.

Cloud security challenges include addressing technology risks (e.g. hypervisor vulnerabilities), risks originating from shared infrastructure (e.g. poor process isolation/data segregation), protection in depth and security at multiple levels (e.g. SaaS application security), resilience and availability (e.g. inability to enforce high-assurance SLAs), data location and mobility (e.g. differences in data protection legislation), information assurance and compliance (e.g. auditing and compliance standards), cloud vendor lock in (e.g. limited service portability), corporate risks (e.g. limited audit capability). Thus, there is still need to engage in research on log and event management in the cloud; monitoring, auditing, compliance and incident management; mitigation of insider threats; cloud-specific forensics tools; business continuity and disaster recovery; enhancing trust and security on cloud infrastructures, trust models in cloud computing, federated identity management for peer assisted cloud or privacy protection as a service in the cloud.

Social networks

Even though social networking enables quick and easy interaction among people, any form of sociality, including the digital one, has always required some voluntary abandonment of privacy. Thus, social networking users must give up some of their private space so as to share it with others. This can lead to privacy drifts such as damaging users' reputation and credibility, security risks (e.g. identity theft) and profiling risks. Accordingly, new kinds of threats have started to emerge (e.g. social engineering, private data misuse by third parties, third party applications, user profiling without the user's consent etc.)

The overarching research challenge in this domain is to safeguard personal freedom and protect personal data, while promoting the free flow and exchange of information. This could be achieved by developing tools and services that on one hand will allow end-users to control their own data, while on the other hand will combine restrictions and rules on data usage with accountability mechanisms. To achieve this, an inter-disciplinary approach should be followed that combines legal, ethical governance and technology aspects.

Mobile

Internet services are increasingly being accessed by mobile clients, such as smartphone and tablets. Thus mobile security is expected to be an increasingly important area of research. The area inherits many of the PC security problems amplified by the specific nature of mobile devices.

Currently, research is conducted towards intrusion detection approaches such as anomaly detection or signature-based mechanisms. These topics remain relevant although new challenges arise:

Personal data management. The huge plethora of sensors that collect information of several types and the need to protect such personal information are two main aspects to be faced. On the one hand we need to protect the user from delivering unwanted personal information, on the other hand we have to balance with the need of applications that need to know information for working properly as in participatory sensing applications for emergency management, where integrity of data should be enforced.

Bring your own device (BYOD). This is a main tend for organizations and it is a very relevant topic, that mixes most of the topics mentioned above.

Repackaging of applications. Several mobile applications available in one market are taken, slightly modified (often with the insertion of malware) and then repackaged and made available in the same market or in others. This is a main vector of infection for end-users as well as of economic damage to application vendors.

IoT

One of the basic principles of the Internet of Things (IoT) paradigm is 'a worldwide network of interconnected entities`. The heterogeneity and diversity of devices and communication protocols make ensuring security of IoT systems a very difficult task. However, taking into account that IoT deployments are used in critical infrastructures, addressing cyber-security threats becomes a major issue. Some of the major challenges in IoT applications are the following:

Data management. Objects are expected to produce and consume data and services. It is therefore essential to assure semantic interoperability between all heterogeneous systems. We assume data and processes are not completely reliable, thus it is necessary to develop efficient collaborative technologies to manage this uncertainty.

Privacy. This is one of the biggest challenges. Other issues include lack of pseudonymity because of static IPv6; remote control of private items by the vendor; establishment of trust 'between things'; how semantic interoperability in Data management create new privacy issues (likability) or issues of Privacy-by-Design

Device level. Security threats and consequent solutions should be considered at the device level and also at the deployment level. For these systems there is a need to ensure confidentiality, integrity, availability, resilience, non-repudiation, authentication, and authorization.

Smart Grid

Smart Grid is an electricity network which composes various users and systems (generators, consumers, providers, etc.) in a cost-efficient way in order to provide a sustainable power system. The system is dynamic and evolving with development of technology. ICT helps Smart Grid to exchange and use various information, including electrical, environmental, and financial data. Thus, the core feature, which makes Smart Grid so attractive from the business perspective, rises a number of security challenges which must be addressed to make this critical infrastructure reliable:

Confidentiality and privacy. Secrecy of the exchanged data must be ensured at all levels of communication inside a Smard Grid. This requires proper cryptographic protocols to be used for data exchange, reliable authorisation mechanisms for all involved devices, safe handling of the information (e.g., safe deletion of data from buffers of devices), etc.

Availability and integrity. Next to proper protocols for data exchanging, Smart Grid must also ensure that the used devices are well protected against tempering and the overall Grid is able to secure the main assets even when some components are compromised. The composite nature of Smart Grid also requires ensuring the agreed SLAs.

Efficient risk management. Since Smart Grid is a dynamic and evolving system, then it should have a similar risk management capabilities. In other words, there is a need to collect security parameters of the system components, aggregate them in an overall picture, analyse the results and make a cost-efficient decision.

Intelligent Transportation Systems

Information technology in automotive systems is gaining importance. The main is reaching efficient and intelligent cars. These new kind of cars increase safety standards, manage fuel/energy efficiently, integrate multimedia capabilities, and raise the comfort of driving and, at the same time, stay economically reasonable to build and maintain.

Among other factors, embedded systems and bus networks, such as these based on the Controller Area Network (CAN), are indispensable for achieving the objectives mentioned above. However, the underlying components consist of hardware and software, which is prone to inspection, infection, and modification by dumping and/or updating the firmware. This holds the potential for a large variety of attacks. Recent research has shown that attacks on automotive systems are feasible and can be launched after gaining physical access or even over-the-air.

As modern cars are equipped with radios for wireless networks (e.g. cellular networks), they can no longer be viewed as independent and sealed off systems. The demand of a higher level of integration with other devices, services and networks, as well as the need of backward compatibility require a well conceived security architecture and a good knowledge of possible attack vectors.

Therefore, security is a very important aspect in the design of automotive systems. They are able to control the fundamental functionality of modern cars as well as safety related tasks which all depend on the security and trustworthiness of the underling technologies.

To address the aforementioned topics there are certain research challenges that can be divided into three major perspectives, which address security issues and requirements from their specific viewpoints. Those three areas are software security, hardware security and security of the underlying mathematical models.

Currently, a number of recent papers cover the growing importance of security in the automotive industry. The University of Washington as well as the University of California have performed research on the topic of automotive security.

Additionally, the European Union prepares new standards that specifies the cooperation of Intelligent Transport Systems (ITS).

Research areas

In this section, we describe several research areas, mainly from a technological perspective, that present interesting research challenges. In particular:

- System security
- Security Engineering
- Security operation management
- Data protection
- Security and big data
- Access control
- Quantitative aspects of security
- Practical and usable security
- Cryptography
- Trust management systems
- Digital freedom
- Network security

Research Area: Systems security

Over the past decade we have been observing a large number of cyber-attacks on the Internet. Starting with the Code Red Worm in 2001, cyber-attacks have demonstrated that they can easily compromise a large number of computers in a short amount of time. Indeed, the early worms compromised tens of thousands of computers in just a few minutes, while lab experiments suggested that very aggressive worms can compromise practically the whole Internet in a matter of seconds. Before these threats became a reality, cyber-attacks swiftly changed in size, volume, and sophistication, submitting their dominance to slow-spreading, stealthy, under-the-radar attacks. Gradually, buffer overflows and code injection attacks, were soon followed by phishing, pharming, and social engineering attacks.

This shift reflects a fundamental change in the profile, motives, and methods of cyber-attackers that happened in a few years. Early attackers were mostly young people, motivated by fun and seeking peer recognition, launched massive, high-profile cyber-attacks in the form of self-replicating computer worms: attacks they could easily brag about. Current cyber-attackers usually serve in the ranks of organized crime or in other illegal organizations. Motivated by financial profits or by political purposes, they usually launch attacks that stay below the radar, are difficult to detect, and exploit the weakest link: the computer user. This change in the motives and profiles has fundamentally changed the problem of computer and network security.

There are a number of threats that systems security must deal with, including malware, botnets, insider threats, targeted attacks, advanced persistent threats, web vulnerabilities, software vulnerabilities, SPAM, malicious hardware, data breaches, social engineering - phishing, passive/active eavesdropping, online behaviour tracking, spoofing – Impersonation, and others. At the same time, systems security will need to deal with emerging environments including on-line social networks, mobile systems, the Internet of Things, and Cloud computing to name a few. In this section we will focus on two flagship problems: (i) how to deal

with malicious software (**malware)**, and (ii) what kinds of threats **online social networks** may pose to their users.

State-of-the-Art

Social Networks: A considerable amount of work has been devoted to protect the identity and support the privacy of users in social network sites. Such systems may include Persona, which uses attribute-based encryption and allows users to dictate policies regarding who may view their information [Bad09], and Safebook, a decentralized and privacy-preserving online social network application [Cut09]. Multiple fake identity (Sybil) attacks on social networks have been used for forwarding spam and malware, out-voting honest users, and manipulating. Identifying spammers in social networks has also received considerable attention. Different methods have been proposed to automatically identify the accounts used by spammers [Str10], and to identify more criminal accounts via the study of social relationships from a number of known malicious accounts [Yan12]. Finally, social links that correspond to interpersonal trust relationships have provided a means to populate white lists of legitimate email senders in Reliable Email (RE) [Gar06], to thwart unwanted communications in OSTRA [Mis08], and to mitigate trust-aware collaborative spam in SocialFilter [Sir11].

Malware: In parallel with the development of cyber-crime into a large underground economy driven by financial gain, malicious software has changed deeply. Originally, malicious software was mostly simple self-propagating code crafted primarily in low-level languages and with limited code reuse. Today, malicious software has turned into an industry that provides the tools that cyber-criminals use to run their business [Rob08, Hay09, Kar05].

Research challenges

- **Malware.** The rise in the number of malware variants continues at a steady pace. Indicatively, McAfee reports a growth in the number of new malware samples of about 8–12 million per quarter for 2012, while as of April 2013 they have more than 128 million malware samples in their database. Symantec reports that in 2012, one in 291 emails contained some form of malware.
  At the same time, the increasing professionalism of cyber-criminals makes defending against sophisticated malware increasingly hard. Once sophisticated tricks of the most skilled virus authors, advanced evasion techniques like code obfuscation, packing, and polymorphism are now the norm in most instances of malicious code. Using polymorphism, the malware is mutated so that each instance acquires a unique byte pattern, thereby making signature extraction for the whole breed infeasible. As the number of new vulnerabilities and malware variants grows at a frantic pace, detection approaches based on traditional string-matching threat signatures, which are employed by most virus scanners and intrusion detection systems, cannot cope with the vast number of new malicious code variants. We need to develop approaches that are able to express the malware of tomorrow, provide accurate signatures for such malware, detect it at line speed, and deploy them at appropriate points that cannot be circumvented by attackers.
- **Social Networks.** The explosive growth rate of social networks has created the first *digital generation*, consisting of people of all ages and backgrounds. People, who are creating their digital counterparts for interacting with other users, may disclose a vast amount of personal data in an attempt to utilize these new services to the fullest. As the on-line social network is a representation of social interaction, it implicitly inherits the trust that may exist between different individuals. However, users may be vulnerable to a series of dangers, ranging from identity theft to monetary loss, and may lack the critical “street-smart” approach that develops over years in the physical world and is passed on from one generation to another. As users tend to show a great amount of trust to online communication and interactions, adversaries may be able to sneak into a victim’s circle of trust through impersonation. As people trust their friends, the cyber-criminal can then perform a

range of attacks that may not be possible, or effective, as a "stranger." We need to develop approaches that would protect users from attacks from strangers, friends, or even the social network itself.

References

[Bad09] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, and D. Starin. Persona: an online social network with user-defined privacy. In Proceedings of the ACM SIGCOMM 2009 conference on Data communication - SIGCOMM '09, pp. 135. ACM Press, 2009.

[Cut09] L. Cutillo, R. Molva, and T. Strufe. Safebook: A privacy-preserving online social network leveraging on real-life trust. IEEE Communications Magazine, Consumer Communications and Networking Series, 47(12):94–101, 2009.

[Gar06] S. Garriss, M. Kaminsky, M. J. Freedman, B. Karp, D. Mazières, and H. Yu. RE: Reliable Email. 3rd Symposium on Networked Systems Design and Implementation, 2006.

[Str10] G. Stringhini, C. Kruegel, and G. Vigna. Detecting Spammers on Social Networks. In Proceedings of the 26th Annual Computer Security Applications Conference, 2010.

[Yan12] C. Yang, R. Harkreader, J. Zhang, S. Shin, and G. Gu. Analyzing spammers' social networks for fun and profit. In Proceedings of the 21st international conference on World Wide Web - WWW '12, pp. 71. ACM Press, 2012.

[Mis08] A. Mislove, A. Post, and P. Druschel. Ostra: Leveraging trust to thwart unwanted communication. In Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation, number i, pp. 15–30, 2008.

[Sir11] M. Sirivianos, K. Kim, and X. Yang. SocialFilter: Introducing social trust to collaborative spam mitigation. 2011 Proceedings IEEE INFOCOM, pp. 2300–2308, Apr. 2011.

[Rob08] R. Roberts. Malware Development Life Cycle. Virus Bulletin Conf., (October), 2008.

[Hay09] M. Hayes, A. Walenstein, and A. Lakhotia. Evaluation of Malware Phylogeny Modelling Systems Using Automated Variant Generation. Journal in Computer Virology, 5(4):335–343, 2009.

[Kar05] M. Karim, A. Walenstein, A. Lakhotia, and L. Parida. Malware Phylogeny Generation Using Permutations of Code. Journal in Computer Virology, 1(1):13–23, 2005.

Research Area: Security Engineering

Building secure services requires that security is considered since the beginning of the Software Development Life Cycle (SDLC) as a holistic process [Joo11, NESSoS]. This includes considering security in the different stages of the SDLC since the process of gathering requirements to the programming and testing phases. Besides that, once the cycle has been completed there should be assurance mechanisms in place that allow us to determine that the resulting service fulfils the functionality it has been built for, and that the whole process complies with the regulations.

State of the art

Traditionally, the inclusion of security in the software engineering process has been misled. Nevertheless, there have been some attempts to consider security in the different phases of the Software Development Life Cycle (SDLC) [SDLC]. During the early phases of the development, security requirements are gathered and possible attacks or threats could be identified. In order to achieve this, enhancements of widely used languages for software specification such as UML have been used. Examples of them are UMLSec [Jur10] and SecureUML [Lor02]. Some authors represent entities and objectives by analysing security from the point of view of the impact on these objectives [Beck13]. In the recent years, researchers and companies realized of the need to approach security in a holistic way in order to build secure services and systems, as for example in [Bas14], where a model-driven approach is used. In the intermediate phases of the development the key is to refine the security concepts included in the requirements phase in order to reflect them into a design architecture. Thus, Mouratidis and Jürjen [Mou10] translate security requirements into design by combining Security Tropos with UMLsec. Other approaches include languages for the specification of authorization and access control policies such as XACML [OAS].

UML statechart tools have been used in the purview of security testing combined with concepts arising from the field of combinatorial mathematics [Boz13], to test for specific interactions between the different components of a system under test (SUT), e.g. a web application. Another approach to the problem of web security testing via casting it to a learning setting has been presented in [Tri13].

At the implementation level the main approaches use secure programming guides or best practices in order to avoid vulnerabilities in the code, although some approaches deal with secure programming languages [Jim02].

One additional problem is how to ensure that the concepts defined in one of the phases remain in subsequent ones, i.e., there are no semantic breaches. Model driven development seems to be a candidate paradigm to solve this problem. Its analogous concept for security is the model driven security paradigm [Bas06]. It consists of building of a general model for security that is being refined up to the point of generating code that implements the required functionality.

Research challenges

- **Security processes.** Security must be addressed from the early stages of the Software Development Life Cycle (SDLC). Processes that include all the mechanisms for designing secure systems should be set in place, providing thus a holistic treatment of security, from the phase of gathering requirements to the phase of implementation, passing through the phase of designing an appropriate architecture for the systems and the implementation. Another aspects to be considered for the life cycle are how security management is treated or users awareness about security. Systems should be adaptive to the needs and security properties have to adapt in real-time. Thus, the inclusion of trust into the processes will help to capture the changes, providing more security.
- **Security testing.** Testing will be part of the process that deals with the implementation phase and it is highly related to assurance. Security testing of web applications still represents a major problem

of software developers and testers alike. In order to reveal vulnerabilities, manual and automatically, software solutions implement different strategies in order to detect certain kinds of inputs that could lead to a security breach. While the first tools require user input from the tester, the second type of tools reduce the necessary amount of user input while still trying to achieve great test case coverage. Both approaches depend on the corresponding test case generation technique that is executed against the system under test. The goal is to find a firm testing pattern that could cover almost every kind of web application. The biggest challenge for security testing is to specify and implement methods in order to detect potential vulnerabilities of the developed system in a never-ending quest against new security threats but also to cover already known ones so that a program is suited against typical attack vectors. These challenges ask for the development of novel techniques in the area of test generation and test execution.

- **Assurance.** Once the services are designed in a secure way we have to ensure that they offer the desired security level. Assurance is a transverse methodology to the process of deriving secure services that should be present along all the phases of the SDLC. Assurance should be considered at different levels: early assurance at the level of requirement, architecture and design and traditional assurance at the implementation level. The latter are techniques related to security testing.
- **Quantitative aspects** of security will support providing assurance. We need to develop metrics and methodologies for assessing the process of assurance for secure services in the FI.
- **Compliance.** Compliance is a property that must be always ensured when designing secure services. From the requirements point of view it is crucial that the new languages that need to be defined for requirements capture notions of compliance. In particular, it is very important to consider legal compliance. The development of secure software should meet compliance regulations such as laws. It would be desirable to develop tools that identify relevant laws to a specific problem in an automated way. Compliance must be ensured as well at the architecture design level. At the implementation level, the verification of the code must be compliant with all the artefacts that are designed in the previous phases of requirements and architecture.

References

[Bas14] D. Basin, M. Clavel, M. Egea, M. A Garca de Dios and C. Dania. A Model-Driven Methodology for Developing Secure Data-Management Applications. IEEE Transactions on Software Engineering, vol 4, issue 4, pp. 324-337, 2014.

[Bas06] D. Basin, J. Doser, and T. Lodderstedt. Model Driven Security: From UML Models to Access Control Infrastructures. In ACM Transactions on Software Engineering Methodologies, 15(1), pp. 39-91, 2006.

[Beck13] K. Beckers, S. Fassbender, M. Heisel, F. Paci and F. Massacci, Combining Goaloriented and Problem-oriented Requirements Engineering Methods. In Proceedings of International Cross Domain Conference and Workshop (CD-ARES), pp. 178-194, LNCS, 8127. Springer, 2013.

[Boz13] J. Bozic, D. E. Simos, and F. Wotawa. Attack pattern-based combinatorial testing. In Proceedings of the 9th International Workshop on Automation of Software Test. ACM, New York, NY, USA, pp. 1-7, 2014.

[Jim02] T. Jim, G. Morrisett, D. Grossman, M. Hicks, J. Cheney, and Y. Wang. Cyclone: A Safe Dialect of C. USENIX Annual Technical Conference, pages 275–288, Monterey, CA, June 2002.

[Jur10] J. Jürjens. Secure Systems Development with UML, Springer-Verlag, 2010.

[Joo11] W. Joosen, J. Lopez, F. Martinelli, F. Massacci: Engineering Secure Future Internet Services. Future Internet Assembly 177-192, 2011.

[NESSoS] NESSoS FP7 NoE D4.3 Part II: Engineering Secure Future Internet Services: A Research Manifesto and Agenda from the NESSoS Community: Final Release available via: http://www.nessos-project.eu/media/deliverables/y3/NESSoS-D4.3-PartII-Roadmap.pdf

[Lod02] T. Lodderstedt, D. A. Basin, and J. Doser. SecureUML: A UML- Based Modeling Language for Model-Driven Security. In Proceedings of the 5th International Conference on The Unified Modeling Language, UML '02, pp. 426– 441, London, UK, UK, Springer-Verlag, 2002.

[Mou10] H. Mouratidis and J. Jurjens. From Goal-Driven Security Requirements Engineering to Secure Design. International Journal of Intelligent System, 25(8):813–840, 2010.

[SDL] Microsoft SDL, http://www.microsoft.com/security/sdl/default.aspx

[OAS] OASIS, XACML Specification 3.0, http://docs.oasis-open.org/xacml/3.0/xacml-3.0-corespec-os-en.pdf

[Tri13] O. Tripp, O. Weisman, and L. Guy. Finding your way in the testing jungle: a learning approach to web security testing. In *Proceedings of the 2013 International Symposium on Software Testing and Analysis*. ACM, New York, NY, USA, pp. 347-357, 2013.

Research Area: Security operation management

It is estimated that, worldwide, more than one million people become victims of cyber-crime every day. The cost of cyber-crime could reach an overall total of USD 388 billion worldwide. A report by Detica (part of BAE plc) that was commissioned by the UK Cabinet Office estimated cyber-crime's annual cost to the UK to be £27bn (about 1.8% of GDP). That report estimated Britain's cyber-crime losses as £3bn by citizens, £3bn by the government and a whopping £21bn by companies. Following some initial disbelief of the size cost estimated, a more thorough analysis by an international team of academics and subject matter experts still reported costs of several hundreds of millions in "genuine" cyber-crime and "transitional" cyber-crime (i.e. crimes whose modus operandi has changed substantially as a result of the move online; for example credit card fraud). Furthermore, advanced persistent threats (APTs) and targeted attacks, cyber-espionage and cyber-terrorism are also on the increase. According to Verizon's research findings, there were 855 targeted attacks incidents and 174 million compromised records among the largest businesses in the USA in 2012. In a recent survey in Europe , 75% of the businesses interview say they have been a concern for some time or that it is an increasing concern, while the majority of respondents believed that their organisation had been the victim of a targeted attack, with 30% reporting a significant business impact.

State of the art

The proliferation of cloud and mobile computing and of social networking, the establishment of e-Government and the emergence of the Internet of Things (IoT) continuous increase of cyber-crime, the emergence of advanced persistent threats (APTs) for corporate and governmental ICT systems, necessitate security research leading to the development of more advanced tools and methodologies for security operations leading to the development of a new and more advanced foundation for future security operation centres. The research challenges cover technological challenges, process and regulatory challenges and challenges relating to people, behaviours and education.

Currently, many organisations are inadequately prepared to deal with intrusions and security incidents: they address the issue only after a serious breach occurs, and when this happens decisions are made in haste limiting the ability to analyse what actually happened and what is likely to follow, to performed and adequately researched identification of the actual source of the incident, the ability to adequately protect sensitive data from leakage, the ability to avoid any disruption to business operations, to adequately support and facilitate forensic investigations. This often leads to evasions remaining undetected and to security operation reactions causing further damage to the system under attack or their dependents.

Modern and future security operations centres and processes should rely on an understanding of the overall security architecture of their realm – be it an enterprise, a government services network or a critical infrastructure such as national network, energy, banking, etc. They need to identify ingress points of attacks and vectors in relation to that security architecture and ensure that the architecture adequately reflects the reality in the field. They also need to cover both physical and logical security and expand visibility and control of their assets both in environments under their control and in partner or outsourced environments including cloud, mobility and IoT.

Technological challenges achieving the above include the size and diversity of data from disparate system, platforms and applications. The lack of unifying security management and governance over numerous point solutions (e.g. anti-virus, firewalls, intrusion detection and prevention, integrity monitoring, log management, identity and access management, etc.) increases the likelihood that some cyber-attacks by malicious activists or criminals or APT-related evasions will succeed by falling between the cracks. The gap between platform and infrastructure protection (typically offered by the provider) and application and data protection (left to the enterprise using the cloud) brings about new opportunities for cyber-criminals to explore. Further to cloud computing, similar gaps in protection, visibility and control between the actors in

complex value networks have emerged in mobility (e.g. with the proliferation of active content and smart enterprise applications) and IoT. Recent regulatory and compliance requirements – such as the emerging European cyber-security directive – also increase the demand for rigorous and accountable incident analysis and reporting for cross-organisational information sharing relating to cyber-incidents. A next generation security operation centre should be at the heart of security functions offering

- **Threat intelligence** about future and forthcoming cyber-attacks or likely evasions,
- Continuously improved and **intelligent prevention** that is sensitive to the context within which the protected assets operate.
- **Continuous monitoring** and near-real time analysis of relevant events complemented with real-time system modelling and situational-/context-awareness capabilities and interfaces that allow exposing (e.g. visually or by using other media) consolidated analytics to the security operations teams.
- **Continuous detection** utilising global threat intelligence and fusing information from multiple sources often supported by additional predictive analytics and uncertainty reasoning tools.
- **Automated, instrumented and instantaneous response** capabilities against threats, remotely exploitable vulnerabilities and security incidents, including the ability to assess the impact of response to the overall system architecture and to roll-back response actions if situational-/ context-awareness indicates a significant change.

Such a security operations centre needs to cooperate closely with CIRT/CERT teams in order to create comprehensive infrastructure for managing security operations. The cooperation between the SOC and CIRT/CERT teams combined the availability of highly trained security operations team members are very important today and will be critical in the future. As control theory experience has taught the scientists and practitioners, the more advance with control system automation and instrumentation and more crucial, and of a higher impact, the contribution and actions of the human operator will be.

The extreme lack of skilled cyber-security operations professionals in the developed world has been recognized in various industry and government publications, while an article published in Computing magazine claims that 21 million more are required in order to properly provide basic protection against threats from hackers and cyber-criminals on the web. Also the recent report from Burning Glass Technologies, which develops technologies designed to match people with jobs, shows that demand for cyber-security professionals over the past six years grew over 2 times faster than demand for other IT jobs. The report is based on a study of job postings for cyber-security professionals placed by U.S. businesses and government agencies over the past six years. In 2013, there were more than 29,700 separate postings for cyber-security-related jobs in a range of industries, including defence, financial services, retail, healthcare and professional services. Cyber-security jobs account for approximately 10% of all IT jobs. The 2013 total is 74% higher than the number of security jobs posted in 2007. By comparison, the number of job postings for all computer jobs grew by about 33% between 2007 and 2013. A very similar increase on the demand for highly skilled cyber-security jobs combined with a lack of necessary skills by job applicants has been witnessed in Europe. For example, the UK has established recently the Cyber-Security Challenge UK , a not-for-profit organization, which is working to encourage talented people with the right mix of skills and know-how to move into the cyber-security profession. A recent survey carried out by the Sans Institute, a sponsor of Cyber-Security Challenge UK, found that while over as 60% of respondents state that demand for recruits is increasing, over 90% are finding it harder to get the people and skills they need. Several EU member state government agencies also report a declining number of students studying computer science at University, which may imply that the skills shortage in the cyber-security sector may continue for up to 20 years.

Research challenges

There are three core areas where training of future cyber-security operation professionals needs to develop – on top of what is currently being addressed by academic and professional organisations:

- The first area includes **training** on methodologies, mathematical models, IT and telecommunications system architectures and solutions, as well as on operational management models. Research in this area also includes design and analysis of optimal Security Operation Centre organisational structures to articulate and attribute roles and responsibilities, and will address security risks associated with the human element in socio-technical systems, as well as the management of the interactions between cyber-security risk management and the overall risk management/continuity plan of an organisation.
- The second area to be addressed is that of **privacy and confidentiality** preserving information exchange and sharing about cyber-security incidents. The focus of this research will be to identify models, system architectures, interaction methods and processes to exchange information on cyber-security incidents of different nature (e.g. technical failures; human mistakes; natural events, malicious attacks) and on threats an vulnerabilities, in order to improve security operations and cyber-security intelligence in complex value chains and ecosystems, encompassing a large number of interconnected players and strongly interlinked information systems.
- The third area uses the above two to further **innovation and advancements** in security systems engineering and cyber-security operations systems management as well as to develop an improved understanding how cyber-attacks evolve including attacks that utilise new generations of malware, advanced evasion techniques and a fusion of computing and social engineering methods including those leveraging trust and influence in social networks.

Research Area: Data Protection

We are living in the era of ubiquitous computing where the decreasing costs for collecting, storing, and processing data combined with the growing number of sensors embedded in different kinds of devices result in an explosion of data that are expected to grow much faster in the coming years. The need of efficiently managing such huge amounts of data has also led to the development of novel computing paradigms (e.g., cloud computing) where data are more and more often stored in remote service providers, which may not be necessarily trusted or trustworthy. These novel paradigms have clearly brought enormous benefits: the availability of a universal access to data; the reduction in power, storage, hardware, and software costs; and the availability of elastic storage and computation services. However, the outsourced data often include sensitive personally identifiable information that is no more under the data owner's control. As a result, the privacy of the data is being put at risk: How do we ensure that our sensitive data remain properly protected? How do we remain in control of who can access our data?

Besides well-known risks of confidentiality and privacy breaches, threats to data remotely stored on a service provider include improper use of information: the service provider could extract, resell, or commercially use substantial parts of a collection of stored data, potentially harming the data owner's market for any product or service that incorporates that collection of data. The protection of data is therefore a key aspect not only in today's digital infrastructure but also for the proper development of future applications in emerging areas such as e-health, bio-banking, ambient intelligence, mobile commerce, financial systems, and so on, which are all characterized by huge amounts of data that need to be shared and processed by different parties.

State of the art

The problem of protecting sensitive information has long been investigated. Most of past work addressed the problem of protecting data in statistical or tabular form [Cir07]. The problem of protecting users' identities and their sensitive information when releasing specific data referred to individuals, called *microdata*, has also received considerable attention, especially following the introduction of *k*-anonymity [Sam01] of which several extensions and variations have been proposed (e.g., [Cap11]). Recently, alternative privacy notions have been proposed such as the differential privacy concept [Dwo06], which has been applied in different application domains (e.g., [Ras09, Xia11]).

Following the emerging scenarios where data and services are stored and managed by external service providers (e.g., in cloud computing and data outsourcing scenarios), research efforts have addressed the problem of protecting data confidentiality and regulating data accesses when data are remotely stored or processed. Particular attention has been devoted to the "honest-but-curious" provider scenarios, where the external service provider, while relied upon for ensuring availability of the data, cannot always be trusted with respect to data confidentiality. The first approaches addressing this data protection issue in an outsourcing scenario typically relied on data encryption and on indexing techniques for enabling query execution (e.g., [Cur11, Cap12, Hac02]). Recent approaches have attempted limiting or departing from encryption whenever possible, working around the idea of splitting data in different fragments stored at different servers or guaranteed to be non linkable (e.g., [Agg05, Cir10, Sam14]). Other solutions have considered the problem of enforcing selective access on outsourced and encrypted data (e.g., [Cap10, Cap13]). These proposals are based on the idea to selectively encrypt data so that users can decrypt only the data they are authorized to access. Selective encryption means that data are encrypted by using different keys and that users can decrypt only data for which they know the corresponding encryption key.

Research challenges

We describe the main research challenges related to the proper protection of data in scenarios where data are stored and processed by external service providers. In the discussion we distinguish between challenges

in protecting *data at rest* (i.e., data that are recorded to a storage device) and *data at use* (i.e., data processed by applications to respond to queries or to make some computations).

Data at rest

- **Data protection techniques.** Whenever we store data at an external service provider, we need guarantees on the fact that their confidentiality, integrity, and availability are properly protected, even to the service provider's eyes. Data protection techniques should be able to satisfy generic privacy constraints corresponding to different privacy needs (e.g., the values assumed by some attributes are considered sensitive and therefore cannot be stored in the clear or the association between values of given attributes is sensitive and should not be released). The proposed solutions should also be robust against possible inferences that can be drawn exploiting data dependencies. Ensuring integrity and availability of data in storage requires providing users and data owners with techniques that allow them to efficiently verify that data have not been improperly modified or tampered with, and that their management at the provider side complies with possible availability constraints specified by the data owner.
- **User empowerment.** When a user, for example, subscribes to a new social networking service or provides some information to access a service, she immediately loses control over the released data. The user's ability in managing her personal information and delete such an information later may then become difficult, if not impossible. Users should therefore be able to specify preferences on which information about them to release – or not release – depending on different factors (e.g., information sensitivity, information recipients or their privacy policies) as well as interaction contexts. Users should also be able to specify restrictions on the secondary usage and dissemination of their own data, allowing them to play a more active role (in contrast to today's passive role of simply declaring acceptance of service provider's practices).

Data at use

- **Fine-grained data access.** Before moving the data to an external service provider, data are often encrypted to protect their confidentiality. Since the storing provider should not have access to the plaintext data, data cannot be decrypted for query execution. Metadata information (indexes) can then be provided for supporting query functionalities. Indexes, however, should be clearly related to the data behind them (to support precise and effective query execution) and, at the same time, should not leak information on such data to observers, including the storing server. The design of inference-free indexes that can be combined with other protection techniques (e.g., fragmentation or access control restrictions) without causing privacy violations are all aspects that still require further investigations.
- **Data computation integrity.** As we move further into the information age, we face many challenges regarding the integrity of computations possibly involving different (and untrusted) data sources. The integrity of computations is a critical issue since the data obtained as a result of a computation are often used to take decisions that may also have an economical impact. Although this problem is not new and many solutions have been proposed, these solutions rely on the presence of trusted components for the verification of the computed results or do not provide a support for complex operations (e.g., many-to-many joins on distributed datasets). An interesting research direction is therefore the design of efficient and effective solutions able to verify the correctness of the results computed through complex operations, also using modern architectures such as MapReduce.
- **Distributed query processing under protection requirements.** The correct definition and management of protection requirements is a crucial point for an effective collaboration and integration of large-scale distributed systems. This problem calls for a solution that must be expressive enough to capture the different data protection needs of the cooperating parties, as well as simple and consistent with current mechanisms for the management of distributed computations, to be seamlessly integrated in current systems.
- **Query privacy.** In several scenarios neither the data nor the requesting users have particular privacy

requirements but what is to be preserved is the privacy of the query itself (e.g., a query searching for treatments for a given illness discloses the fact that the user is interested in the specific illness). It is therefore important to design techniques that enable users to query data while not revealing information about the specific query (i.e., the data the users are looking for) to the server holding the data. Note that effective protection of query confidentiality requires not only protecting confidentiality of individual queries, but also protecting confidentiality of access patterns.

References

[Agg05] G. Aggarwal, et al., "Two Can Keep a Secret: A Distributed Architecture for Secure Database Services," in Proc. of CIDR, Asilomar, CA, January 2005.

[Cur11] C. Curino, et al., "Relational Cloud: A Database Service for the Cloud," in Proc. of CIDR, Jan. 2011.

[Cir10] V. Ciriani, S. De Capitani di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, P. Samarati, "Combining Fragmentation and Encryption to Protect Privacy in Data Storage," in ACM TISSEC, 13(3): 22:1-22:33, July 2010.

[Cir07] V. Ciriani, S. De Capitani di Vimercati, S. Foresti, P. Samarati, "Microdata Protection," in Secure Data Management in Decentralized Systems, Springer-Verlag, 2007.

[Dwo06] C. Dwork, "Differential Privacy," in Proc. of ICALP, Venice, Italy, July 2006.

[Cap13] S. De Capitani di Vimercati, S. Foresti, S. Jajodia, G. Livraga, S. Paraboschi, P. Samarati, "Enforcing Dynamic Write Privileges in Data Outsourcing," in Computers & Security, 39: 47-63, Nov. 2013.

[Cap10] S. De Capitani di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, P. Samarati, "Encryption Policies for Regulating Access to Outsourced Data," in ACM TODS, 35(2):12:1-12:46, Apr. 2010.

[Cap11] S. De Capitani di Vimercati, S. Foresti, G. Livraga, P. Samarati, "Protecting Privacy in Data Release," in Foundations of Security Analysis and Design VI, A. Aldini, R. Gorrieri (eds.), Springer, 2011.

[Cap12] S. De Capitani di Vimercati, S. Foresti, P. Samarati, "Managing and Accessing Data in the Cloud: Privacy Risks and Approaches," in Proc. of CRiSIS, Cork, Ireland, Oct. 2012.

[Hac02] H. Hacigumus, B. Iyer, C. Li, S. Mehrotra, "Executing SQL over Encrypted Data in the Database-Service-Provider Model," in Proc. of ACM SIGMOD, Madison, WI, June 2002.

[Ras09] V. Rastogi, M. Hay, G. Miklau, D. Suciu, "Relationship Privacy: Output Perturbation for Queries with Joins," in Proc. of PODS, Providence, Rhode Island, June-July 2009.

[Sam14] P. Samarati, "Data Security and Privacy in the Cloud, " in Proc. of ISPEC, Fuzhou, China, May 2014.

[Sam01] P. Samarati, "Protecting Respondents' Identities in Microdata Release," in IEEE Transactions on Knowledge and Data Engineering, 13(6), Nov./Dec. 2001.

[Xia11] X. Xiao, G. Wang, J. Gehrke, "Differential Privacy via Wavelet Transforms," in IEEE Transactions on Knowledge and Data Engineering, 23(8):1200-1214, Aug. 2011.

Research Area: Security and Big Data

It is no secret that we are currently situated in the age of Big Data with plenty of challenges and even more promising opportunities. Mostly, Big Data are described as "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" [Gar]. Managing Big Data can become quite complex as data might exist in many heterogeneous data types and formats and stem from a multitude of traditional and non-traditional data sources [Sch12]. Thereby, the processing and analysis of Big Data allows addressing problems having been considered insolvable before or at least not in an economically reasonable way. Considering the enormous benefits the exploitation of Big Data offers, it might seem alluring to ignore potential shortcomings and risks.

The novel circumstances of Big Data analytical practices introduce fundamental questions and increasingly raise concerns regarding data leakage, privacy and surveillance. These topics have especially come into focus since the documents leaked by whistleblower Edward Snowden exposed the wide-ranging data collection activities of NSA and GCHQ. Moreover, there are alarming developments in the modern threat environment.

State of the art.

Since Big Data is a phenomenon that touches multiple domains related to security, it is hard to speak of a state of the art in Security and Big Data. Therefore, the state of the art in a particular domain is briefly discussed in connection with the corresponding research challenge instead of in this dedicated section. Moreover, one has to consider that there are two ways of combining the terms "Security" and "Big Data". This results in research branches discussing *Security for Big Data* as one side of the coin, and in research branches dedicated to the usage of Big Data and related technologies (e.g., NoSQL databases, MapReduce, cloud computing) to enhance *Data-driven Information Security* as the other side of the coin.

Especially for *Security for Big Data*, many research challenges are not entirely new but revived in an aggravated way because of the aforementioned properties of Big Data. *Data-driven Information Security*, in contrast, constitutes a research area that has been established more recently. In the following, the main research challenges related to the combination of security and Big Data are described. Note, however, that the distinction in two major research areas introduced before is mainly done for structuring purposes. In reality, the two concepts cannot be treated strictly on their own because they depend on and support each other.

Research challenges

Security for Big Data

- **Laws and regulations.** Since a large amount of data is somehow related to individuals, the unprecedented opportunities regarding the processing of various kinds of data imply serious privacy concerns. There are laws and regulations addressing different aspects of the data life cycle but most of them are not contemporary anymore and have to be revisited with Big Data in mind. Examples of existing principles that need some adaptation are the definitions of personally identifiable information and informed consent as well as the rigidity of data minimization and data retention. It may even be necessary to introduce new regulations specifically designed for Big Data applications. Furthermore, the principles somehow adaptable to Big Data are mainly limited to privacy concerns. Future work will have to consider other aspects as well. Nevertheless, it has to be kept in mind that legislation is always behind technological innovations.
- **Trust in Big Data – data quality.** As with "ordinary" data processing, working with Big Data requires the establishment of a certain level of trust. In the first instance, the input data have to exhibit an adequate quality for Big Data analytics to be valuable. This includes both the capture of real data and the synthetic generation of realistic and representative data sets. As mentioned, these issues are

already known from traditional data mining but the advent of Big Data makes them even more challenging. Following this, input data of bad quality may lead to untrustworthy inferences and render the final outcomes useless. Consequently, the provenance of result data is another research field connected to trust in Big Data.

- **Trust in Big Data – data processing.** Focusing on the actual processing of the data, trust has to be established in the correct configuration and the accurate functioning of the worker nodes. This issue is exacerbated when the processing activities are moved into the cloud because the organization passes the immediate control over them on to the service providers.
- **Third-party auditing.** Limited trust in the service providers implies a need for ways to monitor their actions as well as to ensure the integrity of the data and the results. Since integrity checks constitute overhead tasks that may make cloud usage less attractive, users should be able to outsource them to an external audit party. Various researchers have used the privacy-preserving public auditing scheme presented in [Wan10] as the basis for new or extended auditing schemes. But they have also been able to reveal several security flaws of it. One important lesson learned from the extensive investigations on flaws is the need for continuous re-assessments of auditing schemes with state-of-the-art attacker models. An additional direction for future work is to address scalability issues of third-party auditing, for example with MapReduce.
- **Security on the technical level.** The storage and processing of Big Data is strongly connected to the developments in the area of NoSQL databases. Since performance has always been the clear priority for them, new security solutions have to keep the database layer thin and must not impair the performance too much. Consequently, most of the research is dedicated to offloading the security components onto frameworks such as Hadoop. Goals of research initiatives proposing changes directly to the core of the Hadoop ecosystem include the development of alternative authentication mechanisms, the implementation of a common audit logging facility, and the support for encryption and key management.

Data-driven Information Security

- **Developments in the modern threat environment.** Because of the economic motivations behind modern attacks, nowadays' adversaries differ significantly from their predecessors in that they are highly skilled, highly motivated, professionally organized and well-funded. In addition, the rising complexity of organizations' IT environments and their openness because of collaboration activities, cloud computing, mobile computing, and BYOD provide several entry points for attackers. These developments have culminated in the advent of the Advanced Persistent Threats (APTs), which are well-researched and targeted against high-value assets. APTs are tailor-made for their specific objectives and operate in a low and slow mode, making them almost impossible to detect with conventional techniques. Consequently, countering APTs constitutes the major motivation for research in data-driven information security.
- **Identity and Access Management (IAM).** With increasingly strong technical capabilities, employees are more and more becoming the weakest spot in organizations' defence structures. Moreover, they are the prevalent targets in the delivery stage of APTs. If organizations do not regularly visit the access rights of their employees and generally have little insight into the actual usage of the privileges, they are likely to face issues like orphaned accounts, privilege creep, and the sharing of accounts with excessive privileges. Promising research efforts in this research area include Identity and Access Intelligence [EMA12] as well as large-scale (visual) role mining [Col12].
- **Information sharing.** Another frequently used method to get access to the internal network of an organization is to exploit zero-day vulnerabilities. Studying these loopholes is a significant challenge for researchers because they cannot be realistically imitated in lab experiments, and the relevant field data is hard to obtain. Thus, it is important to gather as much raw data on actual attacks as

possible and share them among researchers and trusted organizations. This does not only help to examine zero-day vulnerabilities but also to obtain valuable insights on the techniques and strategies of modern adversaries. The most important points to address in order to make more organizations participate in information sharing are to create natural incentives for them and to overcome legal hurdles.

- **Security-as-a-Service.** In recent years, the idea of outsourcing different aspects of information security to service providers has more and more come into focus. This concept can be referred to as "managed services" [Rin14] and is especially relevant to organizations with limited information security resources and expertise (e.g., small and medium enterprises). Since internal knowledge is required at some point to deeply analyse a particular attack, organizations cannot outsource every aspect of information security. But they can get at least those areas as a service where they have considerable shortcomings. With more and more sophisticated technologies deployed in the cloud, Security-as-a-Service is expected to become a valuable market.

References

[Col12] A. Colantonio, R. Di Pietro, A. Ocello and N. V. Verde. Visual Role Mining: A Picture Is Worth a Thousand Roles. IEEE Transactions on Knowledge and Data Engineering 24 (6), pp. 1120-1133, 2012.

[EMA12] EMA, Identity and Access Intelligence: Transforming the Nature of Enterprise Security, Technical report, 2012.

[Gar] Gartner. IT Glossary. Big Data. Online: http://www.gartner.com/it-glossary/big-data/.

[Rin14] T. Ring. Threat Intelligence: Why People Don't Share. Computer Fraud & Security 2014 (3), pp. 5-9, 2014.

[Sch12] M. Schroeck, R. Shockley, J. Smart, D. Romero-Morales and P. Tufano. Analytics: The Real World Use of Big Data. Technical report, IBM Global Business Services, 2012.

[Wan10] C. Wang, Q. Wang, K. Ren and W. Lou. Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing. In Proceedings of the IEEE Conference on Computer Communications (INFOCOM), pp. 1-9, 2010.

Research Area: Access Control

Access control is a critical component of all secure systems, with the combination of authentication and authorization. Both authentication and authorization have received large attention by the research and industrial community, justified by their central role and the many exploited weaknesses observed in real systems. Many recent high-impact security incidents were caused by inadequate access control.

State of the art

The evolution of authentication aims at implementing solutions that are at the same time convenient for users and robust against adversaries who want to circumvent them. The increasing deployment of two-factor-authentication is a clear demonstration of the evolution of the area. Another direction with great potential is the integration with credentials.

Authorizations are also increasing in flexibility and expressivity to manage the complex ICT scenarios, which require managing the activities of a multitude of users, devices, applications, all with heterogeneous security profiles. The simple design criterion of separating a trusted world internal to a company or system from the outside untrusted world is not applicable anymore. Rather, the design should aim at an extensive confinement and isolation of data and applications, with policies that support the system owner in the efficient specification of the restrictions that have to be enforced by the system.

With a classical analogy, efficient brakes allow drivers to go faster. In a similar way, robust security is a necessary component to reap the benefits made available by the evolution of computer and network technology. The analogy is particularly strong looking at the role of access control, which is responsible for keeping the activities executed by computer systems within the boundaries predefined by system and data owners.

Research challenges

- **Flexible Authentication.** Authentication solutions have to continue the progress toward improved user convenience and greater resistance to subversion. The authentication process has to rely on the observation of a number of parameters and rely on redundant ways to verify the identity of the user, adopting simple authentication solutions when the source or profile of requests are consistent with the known user behaviour, switching to alternative authentication solutions when anomalies are observed. Models have to be defined that support the definition of flexible authentication processes, with a richer integration with authorization models. This integration is well represented by attribute-based access control, which uses cryptographic credentials to increases the flexibility of authentication, in such a way that guarantees adequate security.
- **Key management.** A strong strategy to obtain a balance between authentication convenience and security relies on the definition of isolated domains, each one associated with a specific combination of user, data, application and access mode. To enforce protection, each domain is in the end associated with a distinct cryptographic key. The large multitude of keys, associated with different trust assumptions and duration, requires sophisticated key management services. Key management also allows users and systems to derive a multitude of keys from a core of highly trusted secrets. The design of advanced solutions requires a cooperation among the cryptographic and system security community.
- **Access control models for distributed systems.** Data outsourcing is becoming a common occurrence. The encryption and outsourcing of resources creates the need for an evolution of current policy models. In addition to scenarios with a single storage provider, access control models have to support the definition of security requirements for applications that combine resources that are under the control of many independent parties. The secure execution of accesses to data stored in multiple

servers has to consider the relationships between the data. For instance, a citizen can be allowed to access data on the tax declarations that pertain to herself and to her dependents, mixing information stored under different authorities. The construction of applications that integrate data under control of multiple parties requires the definition and support for novel access models, with new constructs that may support the efficient identification of correct access strategies and the detection of potential misuses.

- **Modern mandatory access control.** Mandatory access control models are known to increase the security of operating systems and databases. In the '80s and '90s, the research and industrial communities dedicated a lot of attention to the construction of operating systems and databases with mandatory access control, with limited success at the time. Nowadays, we see more flexible mandatory access control models that increase in importance, becoming a central component in systems that are able to face the threats that arise in today complex scenarios. A clear demonstration of the role of MAC models is represented by SELinux and its integration in Android, SEAndroid. Adaptations of these models can be foreseen for databases and components at other levels. These models represent a great opportunity to realize a flexible confinement of separate application domains.
- **Access policies for information disclosure.** The development of data analysis and data mining solutions has made clear the benefit that can derive from the release of large collections of micro-data, which can support evidence-based investigations on a multitude of topics. These data collections may also offer the opportunity for the retrieval of sensitive information. Access models for security policies are required to make explicit confidentiality constraints that are today implicit, facilitating the construction of systems that are able to offer the benefits of the dissemination of large detailed collections of data and at the same time are guaranteed not to violate the confidentiality of citizens or data owners.

References

[Ard11] C. Ardagna, S. De Capitani di Vimercati, S. Paraboschi, E. Pedrini, P. Samarati, M. Verdicchio, "Expressive and Deployable Access Control in Open Web Service Applications," in IEEE TSC 4(2), pp. 96-109, April-June 2011.

[Agg05] G. Aggarwal, M. Bawa , P. Ganesan, H. Garcia-Molina, K. Kenthapadi, R. Motwani, U. Srivastava, D. Thomas, Y. Xu. Two Can Keep a Secret: A Distributed Architecture for Secure Database Services. In *Proceedings of the CIDR'05*, Asilomar, CA, Jan. 2005.

[Cap07] S. De Capitani di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, P. Samarati. Over-encryption: Management of Access Control Evolution on Outsourced Data. In Proceedings of the VLDB'07, Vienna, Austria, Sept. 2007.

[Cap08a] S. De Capitani di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, P. Samarati. Assessing Query Privileges via Safe and Efficient Permission Composition. In Proceedings of the CCS'08, Alexandria, Virginia, USA, Oct. 2008.

[Cap08b] S. De Capitani di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, P. Samarati. Controlled Information Sharing in Collaborative Distributed Query Processing. In Proceedings of the ICDCS'08, Beijing, China, June 2008.

[Sma13] Stephen Smalley, Robert Craig. Security Enhanced (SE) Android: Bringing Flexible MAC to Android, Proc. NDSS 2013, San Diego, California, Feb. 2013.

Research Area: Quantitative Aspects of Security

It is widely accepted, that perfect security is not achievable. This means, that regardless of the amount of investments there will always be a way to compromise an IT system. Therefore, the main question is not "how to make my system perfectly secure", but something like: "how much security is enough for me?" The answer to this question requires the ability to measure security and compare such measurements. In other words, we need indicators for security and methods for specification of values for these indicators, i.e., security metrics.

State of the art.

Security metrics must satisfy several empirical criteria to be considered as ``good'' and useful metrics [Her07, Jaq11, Swa03, Vau03]. First, a metric should clearly define a *scope* of its application, be a *precise* and *accurate* representation of reality, and have *sound foundation*. The process of metric assessment should be *repeatable* (have the same result after repetition of the assessment), *reproducible* (have the same result if conducted by different evaluators) and *cost-effective*. Last, but not least, a good metric must be *relevant* for decision makers.

Most of the proposed security metrics describe only a very specific aspect of security (e.g., frequency of antivirus updates or number of open firewall ports) [Jaq11, Sto01]. Thus, they are relevant for technical managers to monitor the progress for the considered aspects, but hardly could be relevant for a CISO (Chief Information Security Officer) to justify (additional) investments in security. Thus, currently CISOs lack of a vital instrument to convince the stakeholders that there is a need in security improvements and the ability to share the limited resources efficiently. Often, stakeholders are ready to provide additional investments in security only when a security breach has happened and damage has occurred.

Currently, qualitative metrics, if any at all, are used in practice for this purpose. Qualitative metrics are relatively easy to collect, but hard to aggregate (for a complex system). Furthermore, these metrics are usually subjective, i.e., not reproducible. Finally, the resulting values are imprecise by their nature, and thus, it is difficult to use them for a detailed decision-making process. Therefore, high-level quantitative security metrics are needed.

Many security metrics proposed by researchers are related either with some kind of probability (e.g., the probability of an attack to be successful or the mean time between successful attacks) or cost (cost for an attacker or damage for the system). In other words, most of the metrics relate to risk assessment (risk here is considered as multiplication of frequency of attacks by the possible damage) [Lun11, Sto01, Kra10]. Unsurprisingly, risk assessment received a lot of attention and is required by many guidelines and standards (although, its qualitative version is more popular). Nevertheless, this approach also has a number of problems: the required probabilities are hard to find and quantification of damage is difficult. Furthermore, risk assessment is very cost- and time-consuming [Jaq11].

Research challenges:

- **Quantifying information flow.** There is still a growing interest in automated tools for calculating the flow of information of imperative and concurrent programs. Among the new areas of interest there is differential privacy that seems fairly promising.
- **Selection of good security metrics.** Despite a number of security metrics proposed, it is still unclear how to select the most appropriate metric(s) for decision-making. Many technical metrics often lead to contradictory decisions and singling out the most important one(s) is a challenge.
- **Empirical proofs.** It is hard to compare the proposed security metrics and assessment methods to decide which of them describe security of a system in the best way in practice. This difficulty follows from the confidential nature of the real input data required for the analysis. Only a very few studies

on high-level security metrics contain experimental evidences of validity of the proposed approaches.

- **New metrics.** All proposed metrics so far do not satisfy the requirements for good security metrics completely (apart, maybe of technical metrics applied for a very limited and specific analysis). Therefore, new security metrics are needed to overcome these difficulties.
- **Forecasting security.** Currently, analysis of security is based on past events and the current state of a system. On the other hand, the amount of known vulnerabilities is growing and tools of attackers become more sophisticated. Therefore, what security staff is looking for is prediction for the future in order to know whether the system is well protected for the next months, or what has to be done to make it secure against upcoming threats.
- **Computation of cost and probability of attack.** Cost and probability of an attack are hard to determine precisely. This difficulty impedes usage of quantitative risk assessment. A lot of factors affect the frequency of attacks and the direct impact is not always the most painful for the organization. The damage of reputation, legal costs, loss of trust of partners, etc., are hard to quantify. Moreover, the company also needs to remove attack consequences and may stop its operation or slow down for some period of time.
- **Compositional Risk.** Risk assessment is a time- and cost-consuming process. The problem becomes more grave if we consider a complex system. The idea could be to break the whole assessment process in parts. Thus, every domain manager, which has complete knowledge about his/her domain may focus on his/her part, when a global manager will need simply to aggregate these results.
- **Statistical analysis. Distributions.** Many security metrics rely on statistics to be found (e.g., the probability of an attack). Moreover, some metrics assume that its components are random values and have specific distributions (e.g., breaches are assumed to be modelled with the Poisson process). So far, for general metrics, there is no strong evidence that we can model these components with the specified distributions, even so, it is difficult to find the required parameters of these distributions.
- **Security insurance.** Security insurance is a relatively new type of business, although some companies are on this market for 10 years already. This market is immature if we compare it to other insurance markets (e.g., car, house, health insurances) although it is rapidly growing (thanks to new regulations). Thus, the problem of precise specification of possible damage and probability of attack becomes even more important. Furthermore, more mature schemes of computation of a fair insurance premium are required.

References

[Her07] D. S. Herrmann. Complete Guide to Security and Privacy Metrics. Measuring Regulatory Compliance, Operational Resilience, and ROI. Auerbach Publications, 2007.

[Jaq11] A. Jaquith. Security metrics: replacing fear, uncertainty, and doubt. Addison-Wesley, 2007.

[Lun11] M. S. Lund, B. Solhaug, and K. St_len. Model-Driven Risk Analysis. Springer, 2011.

[Sto01] G. Stoneburner, A. Goguen, and A. Feringa. Risk management guide for information technology systems. Technical Report 800-30, National Institute of Standards and Technology, 2001.

[Swa03] M. Swanson, N. Bartol, J. Sabato, J. Hash, and L. Graffo. Security metrics guide for information technology systems. Technical Report 800-55, National Institute of Standards and Technology, July 2003.

[Vau03] R. B. Vaughn, R. Henning, and A. Siraj. Information assurance measures and metrics - state of practice and proposed taxonomy. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Jan. 2003.

[Kra10] L. Krautsevich, F. Martinelli, and A. Yautsiukhin. Formal approach to security metrics. What does ``more secure'' mean for you? In Proceedings of the 1st International Workshop on Measurability of Security in Software Architectures, ACM Press, 2010.

Research Area: Practical and Usable Security

Human computer interaction is nowadays a big field of research and much effort has been put into improving the practical aspects as well as usability of available technology (eg., [Gup12]).

State of the art

However, the development of security usability has not been equally achieved as security had been mainly developed by security experts and for security experts, so the user has been ascertained as the weakest link of the security chain [Sch00], and thought not to be capable to understand and use security in a correct fashion.

Further, security development has been done with "afterward patches" or in ways that are too complex to be practical and usable. As the main goal of users is not normally to use security [Wes08] but to perform common tasks such as, for instance, web interactions, they are not aware or able to assess what security means are in place in those interactions, and how these can affect their life in practice [Kum10].

More recently, it has been acknowledged that the user *must be* one more component of security solutions and it is now believed that the way security technology is implemented needs changing. But, security experts are just starting to explore this idea while hackers already master the art of social engineering. Attackers only need to find *one* vulnerability to achieve their goal while security experts need to protect the whole system, and this requires finding and fixing *all* vulnerabilities.

Because this research is still in its infancy, there is the need for more organized and stable efforts to focus on the human computer interaction security analysis' research and how the user, as an entity who integrates knowledge, previous social and technological experiences and other human characteristics, who can help researchers to start keeping up with security problems such as: social engineering attacks or even mistakes, bugs and ambiguous interactions, and improve human computer interaction's security usability.

Research challenges

We are all aware and mostly understand physical security, such as the need to lock our home or setup an alarm system. To improve practical and usable security we need to start from the beginning and once and for all stop relying on the "afterward patches" system. The main research challenges to proceed with this research domain are: (a) to better study and understand user's security behaviour; (b) create systematic ways to identify the security problems when interacting with technology; (c) and finally start developing security which can integrate all that knowledge.

- **Understanding user's security behaviour:** There has been some effort in understanding the way users interact with security technology in order to improve its usability as well as decreasing the success rate of social engineering attacks. Related work includes specific security mechanisms/services [Wan14, Bal14]; or the definition of frameworks or conceptual models that try to better understand and integrate all the human ceremonies with technology, in the security analysis [Fer14]. There is the need to better understand how users really interact with technology, if they have security in mind or not, or what really are their objectives, aims, needs, etc. We can start by focusing research on lab experiments and surveys on understanding for what, how and when users interact with security and non-security technology and understand what they take into account (i.e., the diverse contextual, cognitive, personal, and other factors that influence users about their privacy, security and trust) to take secure or insecure decisions. The main goal of this challenge would be to define a model of user's behaviour and interactions with security technology, which includes most human characteristics in terms of diversity, flexibility and human traits.

- **Create a systematic way to identify and test social-engineering attacks or mistakes/ambiguities in the security design, with the help of the user:** once we have a model that closely describes how the user interacts with security, we would need to develop a tool to analyse the security of human computer interactions in a systematic way that can work for and with the help of the user, in order to test, identify and classify social-engineering attacks, mistakes, erroneous or ambiguous interactions. This can be done by creating innovative and closer-to-user ways to detect, alert and even annul the effect of both attacks and inconsistencies.
- **Develop resilient and dependable security for the user:** security technology has been mainly developed by and for security experts, but the fact is that the majority of people are not security experts or, if they are, they can many times risk more. If the first two challenges are fulfilled, we are in better shape to develop security technology as it should have been done some decades ago. If products are to be used by humans, they need to be thought, designed and tested by and for a wide diversity of users, something that nowadays is easy to attain. Only with a diverse, adaptable security technology that can translate human behaviour, needs, intentions, etc., we can start developing more resilient and dependable security, and therefore, more practical and usable.

References

[Gup12] R. Gupta. Human Computer Interaction – A Modern Overview. International Journal on Computer Technology & Applications, vol. 3 (5):1736-1740, 2012.

[Sch00] B. Schneider. Secrets and Lies. John Wiley & Sons, 2000.

[Wes08] R. West, "The Psychology of Security," Communication of the ACM, vol. 51, no. 4, pp. 34–38, April 2008.

[Kum10] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong, "Teaching johnny not to fall for phish," ACM Transactions on Internet Technology, vol. 10, no. 2, pp. 7:1–7:31, Jun. 2010.

[Wan14] Y. Wang, P. Leon, A. Acquisti, L.F. Cranor, A. Forget, N. Sadeh. A Field Trial of Privacy Nudges for Facebook. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, 2014.

[Bal14] R. Balebako, A. Marsh, J. Lin, J. Hong, L. F. Cranor. The Privacy and Security Behaviors of Smartphone App Developers. In Proceedings of the Workshop on Usable Security, 2014.

[Fer14] A. Ferreira, J.-L. Huynen, V. Koenig, G. Lenzini. A Conceptual Framework to Study Socio-Technical Security. Human Aspects of Information Security, Privacy, and Trust. Lecture Notes in Computer Science Volume 8533, pp. 318-329, 2014.

Research Area: Cryptography

Cryptography is the general word used by people and is only the halve of such a scientific activity: the protection of information (data) and resources (hardware, software, communications). This part is the design and the proof of security of cryptographic primitives (e.g. encryption, authentication, integrity, signature, …). The other part is the scientific evaluation and modelling of attacks against such designs: the name is cryptanalysis. The advent of computers and a better use and understanding of mathematics and statistics gave today some advantages to attackers. The possible advent of quantum computers, the bad implementations and the multiple side-channels (passive and active) in realistic implementations are setting security questions about many cryptographic primitives and protocols (not speaking about backdoors and voluntary but hidden weaknesses).

In particular, the breakthroughs in quantum computing such as the Shor's algorithm for prime factorization [Sho94] and Grover's algorithm to invert generic functions [Gro96], which render obsolete the majority of the public-key cryptography as this is used today (e.g. RSA, ECDSA) lightened up the spark for cryptographers to consider alternative methods to construct secure cryptosystems. This effort concentrated on cryptosystems that can allegedly resist attacks mounted by quantum computers, in the sense that the underlying hard primitives which pose as security assumptions, cannot be solved efficiently even in the quantum setting. This alternative field of cryptography is widely known today as **post-quantum cryptography** [Ber09].

State of the art

Encryption- privacy: there are block ciphers (encryption by block of bits) and stream ciphers. Today the design of block ciphers is mainly based on differential and linear cryptanalysis and it is neither giving good lower bounds on the security nor efficient criteria for a strong design.

Authentication-signature: mainly based on RSA, DSA (a variant of Schnorr-El Gamal) and ECDSA (elliptic curves). Problems with padding and (restricted) proofs of security were studied and often solved with benefits for the security.

Hash functions: several new attacks were found and the hash functions MD5 and SHA-1 were demoted for new ones (SHA-2 and later SHA-3). Progress needs to be done for a better provable security.

Key exchange: mainly based on Diffie-Hellman protocols for two entities and generalizations for more entities.

Lightweight cryptography: it is a true living field thanks to the restricted needs of smart cards, RFID and NFC chips.

Database search – cloud computing: the general use of the cloud is pushing the needs for new tools securing data and communications. An active field is the remote questioning of encrypted databases.

Specific applications: voting protocols, auctions, …: a lot of progress done into the direction of provable, secure and practical such applications. However confidence in the entities and implementations is not only a cryptographic problem.

Other points like obfuscation of cryptographic codes, multiparty protocols and provably secure cryptography are living subjects of research.

Cryptanalysis:

- Factorization: slow progress is on the way against RSA, today mainly based on improved implementations and progress in parallel computations. No new idea these last ten years.

- Side-channels: many side-channels are now studied (timing, power, electromagnetism, temperature, sound, …) and are very often very efficient in practical settings. Modelling such side-channels is also on the way with various results.
- Faults: inducting faults during the computation of a cryptographic primitive is a good mean for discovering secret keys. Many results are related to specific implementations. Verification of the result before outputting is not always enough as countermeasure.

A recent algorithmic result on the discrete logarithm problem [Bar14] which concerns a quasi-polynomial algorithm for discrete logarithm in finite fields of small characteristic comes as a reminder that the cryptographic community must be ready to propose alternatives to the number-theoretical systems.

Thankfully, a closer look reveals there is no justification for the leap from "quantum computers destroy RSA and DSA and ECDSA" to "quantum computers destroy cryptography". There are some important classes of cryptosystems which have emerged as alternatives. As pointed out in [Ber09] and in the IST-2002-507932 European Network of Excellence in Cryptology: ECRYPT report on Alternatives to RSA the most promising candidates include: the problem of solving multivariate equations over a finite field, the problem of finding a short vector in a lattice and the problem of decoding a linear code. The latter problems known for being NP-hard serve as the underlying hard problems in the HFE public-key signature scheme [Pat96], the NTRU public-key encryption scheme [Hof98] and the McEliece hidden-Goppa-code public-key encryption scheme [McEl78]. These systems are the cornerstone of multivariate-quadratic-equations cryptography, lattice-based cryptography and code-based cryptography which as mentioned earlier, all together are referred to as post-quantum cryptosystems. There has been significant progress from the time these cryptosystems first proposed; however we limit the state-of-the-art to the latest achievements in the respective subfields of post-quantum cryptography. For example, in the field of code-based cryptography significant progress has been made with regard to structural and decoding attacks. A notorious example is the so-called FOPT algebraic attack [Fau10] which also invalidates the security proof of the CFS signature scheme [Cou01], one of the most well-known schemes in the field of code-based cryptography. A new kind of attack against a variant of McEliece cryptosystem presented recently in [Cou14], with very promising results. Last but not least, efforts have been made to consider alternative hard problems for code-based cryptography in [Sen13]. For multivariate-quadratic-equation cryptography and lattice-based cryptography we refer to the respective chapters of [Bar14], which provide an overview of the latest advancements in these fields.

Research challenges

- **Classical cryptography**: The main challenges are still to find an efficient methodology for provably secure (without restrictions) cryptographic primitives able to fight against many attacks (mathematical, statistical, side-channels, faults, …) and to combine securely these primitives. The problem of correct and trustable implementations (hardware and software) is also of paramount importance. More strong research needs to be devoted to practical cryptography in the real world.

- **Post-quantum cryptography**: In [Ber09] three answers are given - three important reasons that parts of the cryptographic community are already starting to focus attention on post-quantum cryptography (as this also evident from the emergence of the PQCrypto conference series and the building of a dedicated research community, and the identification for RSA alternatives as these are pointed out in the IST-2002-507932 European Network of Excellence in Cryptology: ECRYPT report, http://www.ecrypt.eu.org/ecrypt1/documents/D.AZTEC.2-1.2.pdf):

    - We need time to improve the efficiency of post-quantum cryptography.
    - We need time to build confidence in post-quantum cryptography.
    - We need time to improve the usability of post-quantum cryptography.

In short, we are not yet prepared for the world to switch to post-quantum cryptography. Maybe this preparation is unnecessary and nobody will ever announce the successful construction of a large quantum computer. However, if we don't do anything, and if it suddenly turns out years from now that users *do* need post-quantum cryptography, years of critical research time will have been lost.

References

http://crypto.di.uoa.gr/CRYPTO.SEC/International-Crypto-Security.html

- Workshop, Dagstuhl, June-July 2011 (International View of the State-of-the-Art of Cryptography and Security and its Use in Practice),
- Beijing 2012 (International View of the State-of-the-Art of Cryptography and Security and its Use in Practice II),
- Athens 2013 (International View of the State-of-the-Art of Cryptography and Security and its Use in Practice III),
- International View of the State-of-the-Art of Cryptography and Security and its Use in Practice (IV) December 6, 2013, Bangalore, India.

[Ber09] D.J. Bernstein, J. Buchmann, and E. Dahmen, eds, Post-Quantum Cryptography, Springer. 2009.

[Bar14] R. Barbulescu, P. Gaudry, A. Joux, and E. Thomé. A quasi-polynomial algorithm for discrete logarithm in finite fields of small characteristic. In Proceedings of the EUROCRYPT'13, pp. 1-16, 2013.

[Sho94] P. W. Shor, Algorithms for quantum computation: Discrete logarithms and factoring, Proc. 35nd Annual Symposium on Foundations of Computer Science (Shafi Goldwasser, ed.), IEEE Computer Society Press, pp. 124-134, 1994.

[Pat96] J. Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): two new families of Asymmetric Algorithms. Advances in Cryptology - EUROCRYPT '96 , pp. 33–48, 1996.

[Gro96] L. K. Grover. A fast quantum mechanical algorithm for database search. Proceedings of the Twenty-Eighth Annual Symposium on the Theory of Computing, pp. 212–219, New York, 1996.

[Hof98] J. Hoffstein, J. Pipher, and J. Silverman. NTRU: A ring-based public key cryptosystem. Algorithmic Number Theory. J. Buhler, ed. Springer Berlin Heidelberg. pp. 267–288, 1998.

[McEl78] R.J. McEliece. 1978. A public-key cryptosystem based on algebraic coding theory. DSN Prog. Rep., Jet Prop. Lab., California Inst. Technol., Pasadena, CA., pp. 114–116, Jan. 1978.

[Fau10] J-C. Faugère, A. Otmani, L. Perret, J-P. Tillich. Algebraic Cryptanalysis of McEliece Variants with Compact Keys. In Proceedings of the EUROCRYPT'10: pp. 279-298, 2010.

[Cou01] N. Courtois, M. Finiasz, and N. Sendrier. How to Achieve a McEliece-Based Digital Signature Scheme. In Proceedings of the 7th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology, Colin Boyd (Ed.). Springer-Verlag, London, UK, UK, pp. 157-174, 2001.

[Cou14] A. Couvreur, A. Otmani, J-P. Tillich. Polynomial Time Attack on Wild McEliece over Quadratic Extensions. In Proceedings of the EUROCRYPT'14, pp. 17-39, 2014.

[Sen13] N. Sendrier and D. E. Simos. How easy is code equivalence over Fq? In Proceedings of the 8th International Workshop on Coding and Cryptography, pp. 80–92, 2013.

Research Area: Trust Management Systems

All security models and mechanisms are based on assumptions about the context and environment in which the systems are deployed. Most of these assumptions are explicitly stated, but many important assumptions are implicit in the way that security is modelled and security policies are defined and enforced, e.g. the notion of accountability assumes that all parties are within the jurisdiction of some authority that can resolve conflicts and punish misbehaviour. Such implicit assumptions generally hold in the context for which a system is developed, but they often fail if the context changes or if the system is used in a context that it was not developed for, e.g. when legacy systems are connected through the Internet. Moreover, these implicit assumptions are often embedded in the security model, which makes it difficult to reason about the real security of a system in case they fail.

Trust Management defines a decentralised autonomous approach to security, where all assumptions are made explicit, which makes it possible to reason about them when analysing the security of a system. A complete trust-based security framework (TSF) will only allow an interaction with another agent if the calculated trust in the other agent is sufficiently high given the context in which the interaction will take place. This way, Trust Management systems answer questions about what actions a particular principal is allowed to perform as well as why those actions are permitted.

State of the art

Computational trust and trust management systems have been studied for almost 20 years. Early trust management systems [Bla96, RFC2704] define decentralised credential-based access control mechanisms, which still rely on external authorities to certify attributes. These (first generation) trust management systems allow autonomous specification and decentralised enforcement of security policies. Moreover, they allow dynamic acquisition of credentials to meet the requirements of the access control policies. The policies themselves, however, remain statically defined and cannot evolve to adapt to different or changing environments. Moreover, they work in the context of the existing security model and do not consider the implicit security requirements. The need for better adaptation of policies is met by the second generation of trust management systems, which model the way that people build trust in each other. This supports collaboration with strangers and incorporates third party evidence in the form of reputation and recommendation systems [Jos07, Mas07, Kor09]. There has been a significant drive to model computational trust [Gam88, Mar94, Mck96, Wee01, Car03] and to develop research prototypes that allows empirical evaluation of these trust models in realistic application contexts [Mau96, Cah03, Que06, Zah10]; this drive has primarily been led by European research consortia and institutions.

Trust management systems generally consider two sources types of trust: direct trust (from personal experience) and indirect trust (from reputation and recommendation systems). Direct trust is autonomous and can be established without reliance on trusted third parties, such as authentication- or certification infrastructures. This means that direct trust can be established without knowing the verified identity of the other party [Sei04], so it is possible to improve privacy and reduce the risks of privacy theft through trust management systems [Sei03]. Indirect trust generally relies on input from external entities that an autonomous entity has decided to trust (to some degree). This means that the chain of authority is anchored in the autonomous entity itself, rather than some external entity.

Research challenges

- **Unified Trust Models.** The fundamental concepts of computational trust and trust management systems are now well established and a number of general theoretical and formal models have been proposed by the trust management community. Elements of these models have been evaluated through simulation and empirical work, but an effort to consolidate and unify these general trust

models and to develop accepted evaluation metrics for such general trust models would provide a common platform for the development of future trust management systems.

- **Virtual Anonymity.** The ability to build security frameworks that do not rely on verified identities makes it possible to build systems where persistent pseudonyms or identity escrow systems are used to securely provide services without the service provider knowing the identity of the customer. Exploring the legal and technical scope for providing services to such virtually anonymous users is an important challenge that may profoundly change the way we design ICT systems.
- **Anchors of Trust.** A number of external services, e.g. trusted introducers, trust services and trusted computing, have been proposed to anchor security and trust in highly dynamic systems. The benefits of such services appear obvious, but inclusion of such external trust anchors in a trust management context implies reasoning about the trustworthiness of entities where trust is normally assumed to be absolute. Integration of external trust anchors in an autonomous trust management framework is therefore an important and interesting challenge.
- **Domain Specific Trust Models.** Trust-Based and trust-aware systems have successfully been applied in domains ranging from routing in wireless sensor networks, where interactions are short lived and the number of potential partners is large, to cloud computing, where the duration of interactions are typically longer and the number of potential partners is significantly smaller. It is not obvious whether a parameterised universal trust model or a range of domain specific trust models are best suited to address this diversity, but the development of domain specific trust models will provide important insights into the structures and parameters necessary for trust management systems and it is possible that a unified framework may emerge from a convergence of domain specific trust models.

References

[Bla96] M. Blaze, J. Feigenbaum, J. Lacy. Decentralized Trust Management. In Proceedings of the IEEE Symposium on Security and Privacy, Oakland, U.S.A, 1996.

[Jos07] A. Jøsang, R. Ismail, C. Boyd. A survey of trust and reputation systems for online service provision. In Decision Support Systems, Volume 43, Issue 2, March 2007, pp. 618–644.

[Mas07] P. Massa, P Avesani. Trust-aware recommender systems. In Proceedings of the 2007 ACM conference on Recommender systems, New York, U.S.A. 2007.

[Kor09] T. Korsgaard, C. D. Jensen. Reengineering the Wikipedia for Reputation, in Electronic Notes in Theoretical Computer Science, Vol. 244, pp. 81-94, Elsevier, Aug. 2009.

[Gam88] Diego Gambetta. Can we Trust Trust? in Trust: Making and Breaking Cooperative Relations, 1988.

[Mar94] S. Marsh. Formalising Trust as a Computational Concept. Ph.D. Thesis, University of Stirling, U.K., 1994.

[Mck96] D.H. McKnight, N.L. Chervany. The Meanings of Trust. Technical Report MISRC Working Paper Series 96-04, Management Information Systems Research Center, University of Minnesota, 1996.

[Wee01] S. Weeks. Understanding Trust Management Systems. In Proceedings of the 2001 IEEE Symposium on Security and Privacy, 2001.

[Car03] M Carbone, M Nielsen, V Sassone. A formal model for trust in dynamic networks. In Proceedings of Int. Conf. on Software Engineering and Formal Methods, SEFM 2003.

[Mau96] U. Maurer. Modelling a Public Key Infrastructure. In Proceedings of the 4th European Symposium on Research in Computer Security, Rome, Italy, Sep. 1996.

[Cah03] V. Cahill, B. Shand , E. Gray, N. Dimmock, A. Twigg, J. Bacon, C. English, W. Wagealla, S. Terzis, P. Nixon, C. Bryce, G. M. Serugendo, J-M. Seigneur, M. Carbone, K. Krukow , C. Jensen, Y. Chen, M. Nielsen. Using trust for secure collaboration in uncertain environments. In Proceedings of the IEEE Pervasive Computing, vol. 2, 2003.

[Que06] D. Quercia, S. Hailes, L. Capra. B-trust: Bayesian Trust Framework for Pervasive Computing. In Proceedings of the 4th International Conference on Trust Management (iTrust), Pisa, Italy, May 2006.

[Zah10] T. Zahariadis, P. Trakadas, H. Leligou, P. Karkazis, S. Voliotis. Implementing a Trust-Aware Routing Protocol in Wireless Sensor Nodes. In Proceedings of the 3rd international conference on Developments in e-Systems Engineering 2010, London, U.K., Sep. 2010.

[Sei04] J-M. Seigneur, C. D. Jensen. The Role of Identity in Computational Trust. In Proceedings of the First Workshop on Security and Privacy at the Conference on Pervasive Computing, Vienna, Austria, Apr. 2004.

[Sei03] J-M. Seigneur, C. D. Jensen. Trading Privacy for Trust. In Proceedings of the 2nd International Conference on Trust Management, pp. 93-107, Oxford, UK, Mar. 2004.

Research Area: Digital Freedom

State of the art

Our world is now digital. The revolution that this induces is just beginning with consequences that are much deeper than the industrial revolution. This has consequences in all aspects of human life and interaction with the environment where information dominance becomes properly essential. Security is profoundly impacted and cyber-security becomes a main strategic and sovereignty concern for individuals, economy, societies and nations. These elements are known since at least 40 years but become every day more visible and known of citizens, medias and politicians. Every day examples range from Snowden revelations to the exhibition of movie stars' private nude picture, up to the evidence of first main cyber-wars.

In this context, the societal impact is huge and digital freedom becomes a central attention point to citizens and nations, raising many research questions including typically topics like internet neutrality, e-democracy, online social networks, location-based services and digital health record systems. Let us review some of the research challenges that need now to be solved.

Research challenges

- **E-voting.** E-demography and in particular e-voting is a fundamental constituent of democratic life. Scientific challenges include digital vote protocol design, robustness, simulation of physical voting or definition and studies of appropriate properties. Targeted applications are e-voting asserted processes for small organizational purpose (in associations, companies, ...) to main political elections and a main technological development breakthrough shall consist of providing an open source certified implementation.
- **Identity.** Identity is a central concept of freedom. Without identity, we are freedomless. But identity is also a complex concept in the digital world. Its relation to objects or persons could be complex in particular because of the possible multiplicity of identities.
  The design of the right identity representations allowing in particular to implement anonymization techniques and privacy by design is a difficult and ever evolving research question.
- **Anonymization.** Anonymization is a fundamental process that is needed in many processes that shall be conducted to ensure digital freedom. This is a hard scientific challenge that relies on the ability to certify the robustness of the method over crosscutting information repositories as well as over time. Current main research topics contributing to understand and solve, often under constraints, questions involving anonymity, consist of mastering noise addition, permutation, differential privacy, aggregation, k-anonymity, l-diversity and t-closeness.
  Typical targeted applications are insisting in preservation of privacy, availability of data for research purposes, personal medical file, personal pedagogical file, etc.
- **Privacy.** Privacy by design is the integration of the privacy issues as early as the design phase of a system or application. To become a reality however, its principles must be well defined and supported by methodologies and tools. Among these tools, more emphasis should be put on transparency, in order to provide ways for individuals to understand how their personal data (and, ideally, any data that can be used in a processing with potential effects on them) is collected, generated, managed, transferred, etc. These technologies, which are sometimes called "Transparency Enhancing Technologies", are becoming necessary in a context where information flows are growing dramatically and the data mining and machine learning techniques become more and more powerful.
- **Controlling surveillance.** Digital surveillance is present for the best and the worst in all aspects of our digital life. Its control relies on many different processes going from national and international laws

to underveillance, a concept directly issued from the digital world where surveillance is itself under surveillance.

In this context research on appropriate traceability is needed as well as on robust anonymization techniques.

- **Empowering the user on data.** Every person, connected or not, have digital personal data in the world digital information system. These data appear typically on nation, companies, social security databases as well as in marketing or web companies information systems. How shall we, how can we empower people with respect to these data that directly concern them and on which they currently have no real control?

  Research questions here include traceability of personal data, concepts and means to erase or to make unreadable data under control of individual users. On the accessibility side, capability to present to the users the current data about him present on a given site or on the internet.

  The question of laws design and enforcement, at the international level, is of fundamental interest and shall be investigated involving scientists from informatics and law sciences.

Research Area: Network Security

Advanced cyber-attacks have evolved into an imminent threat and are utilised as part of large scale campaigns for censorship [Lev12], surveillance, espionage and intelligence collection (e.g., PRISM, FOXACID and QUANTUM programs, operated by the NSA), subverting security of cyber-physical systems, e.g., stuxnet, taking down enemy's weapons and defence systems and crippling countries [Les07], and more. New cyber-warfare capabilities are continually discovered as part of the nations cyber-arms race. Such attacks typically utilise vulnerabilities in network protocols, primarily in the routing (esp. inter-domain routing (BGP)) and naming (DNS) systems, which enables remote attackers to intercept traffic or to prevent access to a resource or service, e.g., to downgrade the security of a client system, or for denial of service (DoS) attacks.

Due to the critical functionality that BGP and DNS fulfil, they are frequently exploited for attacks, and although extensively studied, both these systems are still vulnerable. The main vulnerability is that BGP and DNS can be manipulated to force traffic to traverse deliberately selected (malicious) remote hosts. To hijack a traffic destined to a victim network, the attacker can issue false BGP announcements; to redirect a victim to an incorrect (malicious) host the attacker can perform DNS cache poisoning. For instance, these attacks were utilised by the different programs, such as FOXACID, QUANTUM, operated by the NSA, for collection of bulk amounts of data by monitoring users' communication. The state of the art along with the notorious attacks and the recent revelations on the unrestrained surveillance practices, indicate that the Internet's infrastructure is extremely vulnerable.

State of the art

The Internet's routing and naming systems, most notably Border Gateway Protocol (BGP), [RFC4271, RFC1771], and Domain Name System (DNS), [RFC882, RFC1034], are the two core building blocks of the Internet, and are essential for any networked application. Correctness and availability of BGP and DNS are critical to the stability and functionality of the Internet. However, there is a long history of attacks against both.

Inter-Domain Routing (BGP) Security:

The Internet consists of multiple autonomous systems (ASes), owned by organisations, and interconnected by means of routing. To enable connectivity between the different organisations, the networks advertise their address blocks, to the Internet via Border Gateway Protocol (BGP) update messages, [RFC1771, RFC4271]. A BGP message advertises a path to a specific address block hosted by the owner AS. Since BGP does not employ any authentication mechanism to guarantee routing correctness, benign failures or malicious attacks may cause originators of BGP routing announcements to claim address blocks belonging to other networks or to change the routing path to some destination (by adding or removing links), hence rerouting the traffic to incorrect networks. Redirecting the traffic via a different path or network enables censorship, malware and spam distribution and can provide an attacker with man-in-the-middle (MitM) capabilities allowing to intercept, inspect or tamper with, the traffic that traverses it. For instance, such attacks were recently launched to hijack the traffic to Google DNS in Venezuela, and for interception, via Iceland, of the communication exchanged between government offices.

To perform such traffic hijacks attackers craft spoofed BGP announcements that advertise false routes to the real destination IP address blocks, and hence impersonate a legitimate AS, [Mar09]. Unfortunately, BGP announcements cannot be validated, since BGP does not support any authentication mechanism. IP address block hijacking can also expose to Denial of Service (DoS) attacks, by black holing the destination (if the attacker discards the traffic and does not relay it), [Nor04].

To prevent traffic hijacks and guarantee routing correctness, a number of cryptographic defences were proposed, e.g., [Ken00, Ger13] and [RFC6491], that provide authentication of network address blocks.

Unfortunately, none of these security proposals is being (widely) adopted. The only proposal that is seeing some adoption is the Resource Public Key Infrastructure (RPKI), [RFC6491], however, even the RPKI support is available for less than 4% of the networks. The deployment of the proposals requires multiple changes to the routers and protocols. In addition, the proposals typically assume a single trusted root, from which the routers need to establish a chain of trust. Since neither is the root trusted (it is controlled by the US government), nor is it possible to establish a chain of trust to most destinations, the adoption of a secure BGP lingers at best.

Domain Name System (DNS) Security:

The Domain Name System (DNS), [RFC882, RFC1034], enables lookup of Internet services, whose addresses are stored as mappings in the DNS servers. DNS has a long history of cache poisoning attacks, whereby the attackers hijack domains by replacing the records, mapping the service to the authentic address, with spoofed records, mapping the service to attacker's controlled IP address. DNS cache poisoning enables attackers to intercept all the traffic destined to the hijacked domain. Recently, a number of methods were published, [Ger13, Her13a, Her13b, Her13c, Shu14], exposing the DNS servers to practical DNS cache poisoning attacks.

To foil cache poisoning attacks, the IETF standardised DNSSEC, [RFC4034-RFC4035], a cryptographic protection of the DNS records. Although proposed in 1997, DNSSEC is still not widely deployed. Indeed, less than 3% of the DNS resolvers validate DNSSEC records in DNS responses, [Lia13], and less than 1% of the zones are signed.

Recent works [Lia13, Her14] show that cryptographic validation of DNS records results in higher failure rates, increases the amount of exchanged traffic, as well as the latency for clients' applications. In addition, due to their large size, DNSSEC signed DNS responses are frequently exploited in Denial of Service (DoS) attacks to flood victim networks, which in the process also depletes the resources of the abused DNS server. All these further demotivate adoption of DNSSEC.

Research challenges

Securing the core Internet protocols is critical for the correctness, availability and stability of the Internet, of its services, networks and clients.

- **Deployment Challenges.** The first and foremost challenge is identifying the factors impeding adoption of the cryptographic defences for BGP and DNS. This requires investigating the deployment obstacles, infrastructure challenges and implications on other systems (that depend on these protocols).
- **Interoperability with Existing Infrastructure.** Given the obstacles and challenges, the next step is to adjust the proposals to better fit the existing infrastructure.
- **Vulnerable Designs.** Nevertheless, deploying cryptography still does not guarantee security, since a cryptographic scheme constitutes only a small part of the defence, and is often not the target of the attack. Thus an orthogonal challenge is to identify vulnerable designs and to devise countermeasures. Identifying vulnerabilities prior to the standardisation and adoption is extremely important. In particular, the history shows that changing the standards, when the obstacles are discovered, during the deployment phase, causes confusion, extra efforts, additional deployment cycles, failures, and inevitably negative attitude towards the proposed defence among the operational and research community, eventually resulting in impeded adoption. Such evaluation was recently performed on the designs for encryption of DNS traffic, which were shown to be vulnerable to side channel attacks and are non-interoperable with the existing DNS infrastructure, [Shu14].
- **Encryption of All Traffic.** Finally, secure BGP and DNS would prevent most, albeit not all, the attack vectors that currently exist. Unfortunately, eavesdropping on the users' traffic would still be possible.

In particular, most of the DNS servers are located in the US, hence the traffic arriving at these servers is susceptible to surveillance. Furthermore, since the networks in the Internet were not designed according to the geopolitical borders, the paths that the packets traverse between the end points in the Internet, may belong to different countries or to organisations with conflicting interests. In particular, a traffic originating within an EU network and destined to an end point in the EU, can often be routed to traverse networks in other countries, e.g., within the US or China, and stored there for processing. Indeed, the US regulatory framework for intelligence collection specifically allows surveillance of foreign communications conducted on the US soil under the 'Foreign Intelligence Surveillance Act'. Therefore, for protection of the users' privacy, encryption of all the traffic should be supported by the client systems and the Internet services. Unfortunately, according to our analysis of the packets' traces provided by CAIDA, less than 6% of the Internet traffic is encrypted.

References

[Lev12] P. Levis. The Collateral Damage of Internet Censorship by DNS Injection. ACM SIGCOMM CCR 42(3), 2012.

[Les07] M. Lesk. The New Front Line: Estonia Under Cyberassault. IEEE Security & Privacy 5(4):76–79, 2007.

[Mar09] C.D. Marsan. Six worst internet routing attacks, 2009.

[Nor04] O. Nordström, C. Dovrolis. Beware of bgp attacks. ACM SIGCOMM Computer Communication Review, 2004.

[Ken00] S. Kent, C. Lynn, K. Seo. Secure border gateway protocol (S-BGP). IEEE Journal on Selected Areas in Communications, 18(4):582–592, 2000.

[Ger13] J. Gersch, D. Massey. ROVER: Route Origin Verification Using DNS. In Proceedings of the 2013 22nd International Conference on Computer Communications and Networks, IEEE, pp. 1–9, 2013.

[Her12] A. Herzberg, H. Shulman. Security of patched DNS. In: Computer Security - ESORICS 2012 - 17th European Symposium on Research in Computer Security, Pisa, Italy, September 10-12, 2012. Proceedings. (2012) 271–288

[Her13a] A. Herzberg, H. Shulman. Fragmentation Considered Poisonous: or one-domain-to-rule-them-all.org. In Proceedings of the IEEE CNS 2013. The Conference on Communications and Network Security, Washington, D.C., U.S., IEEE, 2013.

[Her13b]A. Herzberg, H. Shulman. Vulnerable delegation of DNS resolution. In Proceedings of the 2013 - 18th European Symposium on Research in Computer Security, Egham, UK, 2013.

[Her13c] A. Herzberg, A. Shulman. Socket Overloading for Fun and Cache Poisoning. In Proceedings of the ACM Annual Computer Security Applications Conference, New Orleans, Louisiana, U.S., Dec. 2013.

[Shu14] H. Shulman, M. Waidner. Fragmentation Considered Leaking: Port Inference for DNS Poisoning. In Proceedings of the Applied Cryptography and Network Security, Lausanne, Switzerland, Springer, 2014.

[Lia13] W. Lian, E. Rescorla, H. Shacham, S. Savage. Measuring the Practical Impact of DNSSEC Deployment. In Proceedings of the USENIX Security, 2013.

[Her14] A. Herzberg, H. Shulman. Retrofitting Security into Network Protocols: The Case of DNSSEC. Internet Computing, IEEE 18(1):66–71, 2014.

[Shu14] H. Shulman. Pretty Bad Privacy: Pitfalls of DNS Encryption. In Proceedings of the 13th annual ACM workshop on Privacy in the electronic society, 2014.

Other aspects

In this section, we highlight some aspects that emerged from the discussions and we think we worth noticing, in particular:

- Escaping the legal trap;
- Crash Commission" for cyber?
- Improving Awareness.

Other aspects: Escaping the Legal Trap

Tim Berners-Lee coined the term of Web science[1]. By this he understood that the Web has not only a technical dimension, but also a social dimension and that both are entangled. He also understood by it that changes in the technical aspects will trigger social changes. At the same time, social paradigms will influence how we use or abuse a technical system of large scale like the Web. It is this aspect and this interface between the technical and the social dimensions that can enrich the horizon ERCIM may explore in scientific research. We may discover and understand the issues better and we may find solutions that take both aspects into account.

Privacy is a core domain of Web science, because it was born in the social and legal scientific community and has large technical connotations. Including Security does not look as obvious. But IT -Security has many social aspects. There is hacking, but there is also social hacking. There is data breach and there are laws regulating data breach notifications. The current trend in Security tries to find behavioural patterns to detect attacks, classify and thus further intrusion detection. This in turn has Privacy implications as the pattern matching exercise requires a complete picture of the network traffic involved and is thus not distinguishable from pervasive monitoring. The application and implementation of certain Security management standards will decide on the liability of a service. Now if on the one side, the Security management imperatives proscribe an extensive data collection practise in order to allow for meaningful security pattern matching and Privacy imperatives proscribe data minimisation, the rule system has created conflicting imperative goals, a legal trap.

Resolving the conflict between two or more important goals set by a rule system is not trivial. Today, often only one side has to resolve the issue, mostly the social or legal side that has set those conflicting goals. It can be seen as an exception when the legal debate is technically informed and takes constraints of the technology stack into account. Often, a slight change in the protocol would have avoided a major legal issue. The rule is either mutual ignorance or a battle between both sides over the last word. The last word will then determine whether the technical side or the legal side has to live with the disadvantages of a solution that only maximises the requirements of the other side, respectively. This ranges from copyright protection over the patent system to the questions of Privacy and Security.

The counterculture tries to build bridges between those setting social goals and those implementing things with real world consequences. The bridge is enshrined in terms like "Privacy by design" and "Security by design" that express more of a political vision and can hardly be used to find concrete recipes for action to escape the legal trap. The difficulty stems from the fact that those experiencing the legal trap are not legally savvy and those creating the legal trap will not recognise it as such because they lack the technical understanding. Escaping the legal trap thus has three dimensions: The technical dimension, the legal dimension and the communication dimension. Not taking the third one into account bears the risk that we continue what the French characterise as "dialogue de sourds" the dialogue of the deaf.

1 http://www.w3.org/2009/Talks/1109-websci-tbl/

Now this was all abstract and theoretic. The litmus test will be to find a concrete example where all the above proves to be true. The recent decision of the European Court of Justice could deliver such an example. In Spain vs. Google[2], the ECJ decided that a data subject can turn against the information provider or the search engine to remove personal information and links to that personal information[3].

A local journal in Cataluña, La Vanguardia, has a PDF archive of its printed pages. On page 23 of La Vanguardia from 19 January 1998, there is a right column with small print announcements. One of them being for public sale of an apartment because of some debt to social security, including the name of the owner of the apartment. Mr González turned to the Agencia Española de Protección de Datos (Spanish Data Protection Agency; 'the AEPD') and requested La Vanguardia to remove the pages or alter them to avoid their indexing. He also requested that Google would not show the results anymore.

To cut a long story short, AEPD decided that there was still a legitimate reason for the content to be online and the ECJ had to decide whether M. Gonzalez would be entitled to ask Google to remove links to the announcement of public sale of his apartment. Google challenged this in court. The court put the questions about EU law in front of the ECJ and the ECJ decided. Now Google has to erase links. But what seemed obvious in the beginning raises more and more questions. Those questions certainly touch on the legal aspects of data protection law and free speech, but they also touch on the technical questions concerning the Web.

The information is still there, but cannot be found anymore by Google. Despite the fact that other search engines still find the content, we assume that in the name of equal treatment of all in front of the law, they should also remove the links to the content containing the name. But what is the purpose of a digital archive that cannot be searched? What is the difference between what we now call an archive and a web server that serves historical content and is indexed by search engines as every other web server? How can we limit the outreach of public and online available archives? Deciding that Google has to erase a link does not answer those questions. To the contrary, the discussion is now about search engines and not about archives. We can already see that the current black & white rules from data protection are not really fit for purpose and do not address the technical challenges. E.g. that Google must remember all the forgotten content to avoid a re-indexing. That means we have at Google a concentration and profile of people who have some information online that embarrasses them. Now Google is the biggest holder of profiles already. Data protection was initially done to prevent too much power in one single hand. But now data protection regulation forces the building of profiles in the hand of an actor that has already too much power. So the technical constraint of re-indexing is forcing the actors to do exactly the contrary of the overall goal of the data protection law. If we take into account that Google can be forced to hand over that list to the NSA, this gets yet another dimension. By forcing the centralised data management control within Google, the legal system has exactly followed the legal system requirements. But what would technologists have recommended to the judges?

In fact, there is technology that could help to decentralise the information about the information that should not be indexed. The search engines follow some rules they created themselves years ago. If a web server has a file named robots.txt in the root of the file-tree served, the search engine's crawler will read it and follow the directives written in there. This would force the data subject to address the publisher of the information and not the intermediary, here the search engine. The information about removed links is not in one hand anymore as it is distributed over all concerned content providers. Technically, this solution is superior. But robots.txt is still very basic. The format is not powerful enough to address parts of a page or things in a dynamic environment. We could develop and promote an extension to this format that allows to take the

2 ECJ C-131/12 http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62012CJ0131&rid=2

3 Legal analysis of the court decision by the author: http://www.internet-law.de/2014/05/the-ecj-is-right-the-result-is-wrong.html

Open Web Platform into account. This needs multi-disciplinary research into what is needed from the legal side and a creative technical applied search for scalable solutions.

Now have we passed the litmus test with the ECJ example? The legal dimension is addressed by the fact that data protection law has set some goals. The technical dimension is addressed as the obvious legal solution leads to more problems instead of leading to a solution. Now comes the communication. And there research should be organised in such a way that puts researchers in engineering and legal aspects into a single fishbowl where they cannot escape each other. By forcing them to publish common papers, they have to take responsibility for each other. So yes, the example passed the test.

As a result, such a setup can find the traps before they affect people and offer a solution before a court makes uninformed decisions. Finding those traps will help close the gap between technology and the legal system and thus lead to a more credible legal system that has a much higher accepted authority.

Challenges

As we are looking into the future, we have to identify some core topics for our research resulting from the above. Suggestions are:

- How to **resolve the tension between security logging and legal requirements for privacy**. As we know, the ECJ has invalidated the data retention directive 2006/24/EC[4] with decision C-293/12[5]. This raises many questions on how to treat the ambient log files and how to find solutions that can help find intrusions and follow criminals but still preserve the privacy of all others.
- How to allow personalisation to **help** us **navigate the information jungle without putting too much power in one hand**. This involves decentralisation of profile information and distributed control over such information.
- How to **make data protection law implementable without losing functionality**. This may lead to concrete suggestions for changes in the legal landscape but may also enlighten the understanding of what "Privacy by design" really means.

4 http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0024

5 http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62012CA0293

Other aspects: “Crash Commission” for cyber?

Critical infrastructure is increasingly controlled through ICT, from transport - Air Traffic Control (ATC) and Automatic Train Control Systems (ATCS) - to utilities, such as electricity (Smart Grids). The importance to society of such critical infrastructure means that critical infrastructure providers have developed a strong safety culture to identify problems that may cause disruption to services and ways to improve the robustness of the critical infrastructure. In many cases, this safety culture includes collaboration with an independent agency to establish the cause of all disruptions or "near miss" incidents, such agencies are commonly known as crash commissions or accident investigation commissions. In Europe, the European Aviation Safety Agency (EASA) is responsible for aviation security and the European Maritime Safety Agency (EMSA) is responsible of security at sea. All infrastructure providers are required to report all incidents to these agencies, which conduct a thorough investigation in order to establish causes and make rules and recommendations about how similar incidents can be avoided in the future. While these agencies may investigate the role of failing ICT in reported incidents within their respective domain, there is no agency to investigate ICT failures in general.

The increasing importance of ICT in all aspects of modern society means that failing ICT systems may impact all aspects of our daily life, from the inability to pay our bills through online banking or contacting government services online to large databases of our personal information being leaked on the Internet. The "safety culture" of most ICT system developers, outside the domain of critical infrastructure, has historically been rather poor and there are frequent reports of security failures in ICT systems that are widely distributed and generally relied upon. It is our belief that the poor security record of ICT can be improved significantly by following the example of critical (transport) infrastructure providers, i.e. through the establishment of an independent crash commission for ICT. It should be made mandatory for citizens, government agencies and industry to report both security failures and “near misses” to this crash commission and to make all information available to the crash commission that is necessary for it to complete its work. It would be the task of the crash commission to investigate all incidents in order to establish all the causes for the incident and the relative importance of each individual cause in the resulting incident. The crash commission should make its findings public in a form that allows all the parties to learn from the mistakes that caused the incident without further compromising the security and operational integrity of the entity reporting the incident.

Challenges

- **Acceptability of the reporting obligation.** It is important to ensure a broad accept of the reporting obligation from citizens, government agencies and industry, because publication of security failures may affect an entity’s reputation and result in a loss of confidence from the partners that the entity normally collaborates with.
- **Establishing reporting requirements.** In order to facilitate the work of the crash commission, it is important to identify the types of information that should be reported to the crash commission and define suitable data formats for the reporting of information to the crash commission.
- **Crash commission reports.** The reports from the crash commissions should provide sufficient information for other organisations to draw the right conclusions and avoid unsafe practises and technologies.
- **Automatic reporting infrastructures.** While it may be possible to compel government agencies and industry to report incidents through legislation, many ordinary citizens will have neither the ability nor the inclination to report every security related incident, so compulsory automatic reporting, e.g. from personal firewall or anti-virus products, may be considered. As this may ultimately be seen as a violation of privacy, it is important to balance the protection of privacy against the possible public

good and to define appropriately anonymised reporting structures for such an automatically submitted reports.

Other aspects: Improving Awareness

ICT are now indispensable to the regular functioning of modern society. There is an increasingly dependence on the regular operation of information systems, communications infrastructure and mechanisms to implement the legal guarantees of rights and freedoms of citizens. Threats to their availability, integrity and confidentiality can result in disastrous occurrences for the normal functioning of most institutions. However, while the adoption of new technologies is high, most users remain unaware of their exposure to risks from security flaws. Traditional approaches to tackle this issue are based on the adoption of security policies, short security courses in induction training for staff and new regulations. However, given the criticality of this issue, new approaches should be investigated in order to increase the level of awareness.

Challenges

- **Raising awareness through serious game(s).** The term serious game is used for game-based situations used for non-leisure purposes or serious applications such as learning and training. The use of serious games for learning or training is a trend, which has increased lately due to the relative availability and ease of use of the Internet and increasing broadband connectivity [Mal87]. Serious games not only open up the possibility of defining learning game-based scenarios but also of enabling collaboration among players that might lead to better learning outcomes [Tud92].
  The goal of this research challenge is to explore the effectiveness of serious game(s), particularly their impact in raising awareness of the issue of information security and privacy. The game should also serve as a platform to formally define and study concepts of awareness, what it means and how to measure it to assess societal impact of the proposed game(s).
- **Raising awareness through massive open online course.** The term massive open online course (MOOC) is used for online course(s), in addition to traditional course materials such as videos, readings, and problem sets, MOOCs provide interactive user forums that help build a community for students and professors.
  The goal of this research challenge is to explore the effectiveness of MOOCs, particularly their impact in raising awareness of the issue of information security and privacy, as they are mainly targeted to an unlimited audience with open access via the web.

References

[Mal87] Malone, T. and M. Lepper. Making learning fun. Aptitude, Learning and Instruction: Conative and Affective Process Analyses. R. Snow and M. Farr, Lawrence Erlbaum: 223-253, 1987.

[Tud92] J. R. H. Tudge. Processes and consequences of peer collaboration: A vygotskian analysis. Child development 63: 1364 – 1379, 1998.

Conclusion

The research topics identified in this document represent interesting challenges to be addressed by ERCIM partners, that have a strong expertise and competence in the research and application domains mentioned above.

It is clear that security and privacy are two main concerns of modern societies that dramatically and interestingly do not witness a reduction of interest. While citizens are often surprised by security failures and privacy concerns (see the recent Snowden cases), one could note that many positive things happen and are possible due to the technologies….

ERCIM is willing to continue to perform research activities in these challenges fields by continuing to provide innovative research and tools for the security and privacy of the European citizens.

The interested reader can find further information on the ERCIM activities in security and privacy at the following URL: http://www.iit.cnr.it/STM-WG/ of the ERCIM Security and Trust Management WG (STM).

Contributors

Luís Antunes

Michele Bezzi

Sabrina de Capitani di Vimercate

Carmen Fernandez-Gago

Theo Dimitrakos

Sotiris Ioannidis

Christian D. Jensen

Sokratis Katsikas

Claude Kirchner

Javier Lopez

Volkmar Lotz

Ana Margarida

Evangelos Markatos

Fabio Martinelli

Sjouke Mauw

Stefano Paraboschi

Günther Pernul

Christian Richthammer

Peter Ryan

Jean-jacques Quisquater

Pierangela Samarati

Haya Shulman

Dimitrios Simos

Michael Waidner

Edgar Weippl

Rigo Wenning

Artsiom Yautsiukhin