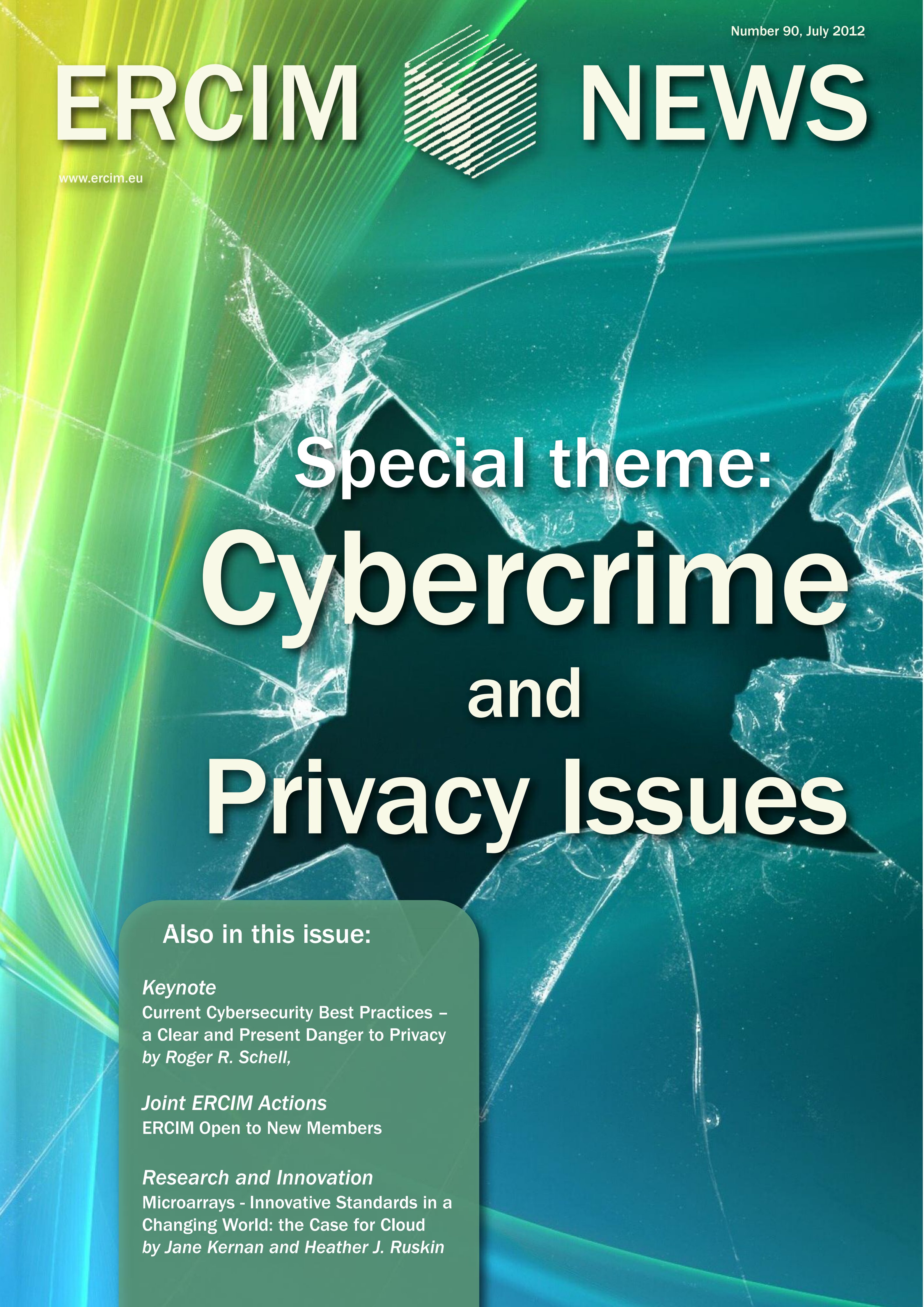


Number 90, July 2012

ERCIM NEWS

www.ercim.eu

Special theme: Cybercrime and Privacy Issues

Also in this issue:

Keynote
Current Cybersecurity Best Practices – a Clear and Present Danger to Privacy
by Roger R. Schell,

Joint ERCIM Actions
ERCIM Open to New Members

Research and Innovation
Microarrays - Innovative Standards in a Changing World: the Case for Cloud
by Jane Kernan and Heather J. Ruskin

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 8,500 copies. The printed version of ERCIM News has a production cost of €8 per copy. Subscription is currently available free of charge.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
Tel: +33 4 9238 5010, E-mail: contact@ercim.eu
Director: Jérôme Chailloux
ISSN 0926-4981

Contributions
Contributions must be submitted to the local editor of your country

Copyright Notice

Advertising
For current advertising rates and conditions, see
http://ercim-news.ercim.eu/ or contact peter.kunz@ercim.eu

ERCIM News online edition
The online edition is published at http://ercim-news.ercim.eu/

Subscription
Subscribe to ERCIM News by sending email to
en-subscriptions@ercim.eu or by filling out the form at the
ERCIM News website: http://ercim-news.ercim.eu/

Next issue
October 2012, Special theme:
What is computation? Alan Turing's Legacy

Roger R. Schell, President of ÆSec, founding Deputy Director of the (now) US National Computer Security Center. He is considered as the "father" of the Trusted Computer System Evaluation Criteria (the famous "Orange Book")

Current Cybersecurity Best Practices – a Clear and Present Danger to Privacy

Not only is the effectiveness of current cybersecurity "best practices" limited, but also they enable and encourage activities inimical to privacy. Their root paradigm is a flawed reactive one appropriately described as "penetrate and patch". Vigorous promotion encourages reliance on these flimsy best practices as a primary defense for private information. Furthermore, this paradigm is increasingly used to justify needlessly intrusive monitoring and surveillance of private information. But even worse in the long term, this misplaced reliance stifles introduction of proven and mature technology that can dramatically reduce the cyber risks to privacy.

Threat of software subversion is dire risk to privacy

Today much of the private information in the world is stored on a computer somewhere. With Internet connectivity nearly ubiquitous, it is the exception – rather than the rule – for such computers to be physically/electrically isolated, i.e., separated by an "air gap". So, protection for privacy is no better that the cybersecurity best practices defenses employed, and their evident weakness attracts cybercriminals. Billions of dollars of damage occur each year, including identity theft with massive exposure of personal data. Clearly weak cybersecurity defenses create a serious risk to privacy.

Juan Caballero's article in this issue notes that "At the core of most cybercrime operations is the attacker's ability to install malware on Internet-connected computers without the owner's informed consent." U.S. Navy research demonstrates that an artifice of six lines of code can lay bare control of a commercial operating system. The Stuxnet, DuQu and Flame software subversions have recently been detailed, and a senior researcher wrote, "Put simply, attacks like these work." I made the point myself in a 1979 article on "Computer Security: the Achilles' heel of the electronic Air Force?" where I characterized subversion as the technique of choice for professional attackers.

Best practices are not well aligned with the threat

The general response seems primarily to be a concerted push for the use of best practices, with a heavy emphasis on monitoring techniques like antivirus products and intrusion detection. For example several Silicon Valley luminaries recently presented a program with an explicit goal "To promote the use of best practices for providing security assurance". In the litigious U.S. there have even been legislative proposals to reward those who use best practices with "immunity" to lawsuits.

Yet this fails to align with the software subversion threat. A major antivirus vendor recently said, "The truth is, consumer-grade antivirus products can't protect against targeted malware." A FBI senior recently concluded that the status quo is "unsustainable in that you never get ahead, never become secure, never have a reasonable expectation of privacy or security". Similarly, an IBM keynote presenter said, "As new threats arise, we put new products in place. This is an arms race we cannot win."

But, it is even more insidious that governments use the infirm protection of best practices as an excuse for overreaching surveillance to capture and disseminate identifiable information without a willing and knowing grant of access. They falsely imply that only increased surveillance is effective. In fact, dealing with software subversion by a determined competent adversary is more intractable than scanning a lot of Internet traffic, as Flame and StuxNet amply demonstrate.

Proven verifiable protection languishes

In contrast, the security kernel is a proven and mature technology developed in the 1970s and 1980s. Rather than reactive, security is designed in to be "effective against most internal attacks – including some that many designers never considered". The technology was successfully applied to a number of military and commercial trusted computer platforms, primarily in North America and Europe. It was my privilege to lead some of the best minds in the world systematically codifying this experience as the "Class A1" verifiable protection in the Trusted Computer System Evaluation Criteria (TCSEC). An equivalent technical standard promulgated in Europe was known as ITSEC F-B3, E6.

Although no security is perfect, this criterion was distinguished by "substantially dealing with the problems of subversion of security mechanism". In other words, a powerful system-level solution aligned with the threat in just the way glaringly missing from current cybersecurity best practices. Unfortunately, at that time addressing this threat was not a market priority.

Although still commercially available, the technology has fallen into disuse in the face of the expedience of the reactive paradigm. It is particularly disappointing that now at the very time ubiquitous Internet connectivity makes privacy really, really interesting, educators and industry leaders have mostly stopped even teaching that it's possible. But today's researchers have one of those rare tipping point opportunities to lead the way to recovery from the clear and present danger to privacy by aggressively leveraging that proven "Class A1" security kernel technology.

Roger R. Schell

KEYNOTE

JOINT ERCIM ACTIONS

SPECIAL THEME

This special theme section on "Cybercrime and Privacy Issues" has been coordinated by Jean-Jacques Quisquater Université catholique de Louvain, Solange Ghernaouti-Hélie, University of Lausanne, Jens Tölle and Peter Martini, Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE

Introduction to the special theme

Invited article

Invited article

RESEARCH AND INNOVATION

This section features news about research activities and innovative developments from European research institutes

EVENTS, BOOKS, IN BRIEF

ERCIM Open to New Members

ERCIM, a consortium of leading research institutions, is opening its doors to new members. The organization, which focuses on information and communication science and technology (ICST) and related areas of mathematics, has a successful track record of promoting ICT research and cooperation in Europe and beyond. Membership was previously restricted to one member per country, but that limit is now lifted to allow applications from more top level research institutions and university departments in ICT from each country.

ERCIM aims to foster collaboration within the European ICT research community and to increase cooperation with industry. It currently has 20 centres of excellence across Europe and is internationally recognized as a major representative organization in its field. ERCIM provides access to all major ICT research groups in Europe and has established an extensive program of working groups, publications, fellowships and prizes. It also hosts the European branch of the World Wide Web Consortium (W3C).

Activities

ERCIM has an excellent track record of successful initiatives to promote ICT research and cooperation in Europe and beyond. These include:

- Working Groups: ERCIM runs Working Groups on topics ranging from computing and statistics to software evolution, preparing the way for top level research in new domains.
- The "Alain Bensoussan" Fellowship Program which has attracted more than 300 post docs since its inception, currently supported by the European Commission FP7 Marie Curie Actions .
- ERCIM News, ERCIM's quarterly magazine with a circulation of 9000 copies, as well as free on-line access, provides a forum for the exchange of information between the institutes and also with the wider scientific community.
- The prestigious Cor Baayen Award, presented each year to a promising young researcher in computer science and applied mathematics.
- Cooperation with professional bodies, specifically with the European Mathematical Society (EMS), the European Telecommunications Standards Institute (ETSI), the European Science Foundation (ESF) and the European Forum for ICST (EFICST).
- Strategic advice for European and non-European bodies, realised through studies, strategic workshops and leadership of or participation in expert groups.
- Research project management support: ERCIM has successfully applied for and managed numerous leading research projects in the range of European Framework Programs.

Benefits of ERCIM membership

ERCIM is a European-wide network internationally recognized as a representative organisation in its field so members can benefit from easy access to all major ICT research groups in Europe. Members can take part in all ERCIM activities including research projects, Working Groups or in the PhD fellowship programme supported by the European Union. They also benefit from ERCIM's privileged partnership with standardisation bodies such as W3C and ETSI.

How to become a member

Prospective members must be outstanding research institutions within their country. Applications will be reviewed by an internal board and might include an on-site visit. Membership is renewable as long as the criteria for excellence in research and active participation in the ERCIM community, cooperating for excellence, are met.

Members must be head-quartered in Europe, where Europe is defined as the European Union Members countries and the European Free Trade Association (EFTA) Member countries. In exceptional circumstances the General Assembly can admit a member not fulfilling this criterion.

For further information about how to join ERCIM AISBL, please contact Domenico Laforenza, ERCIM vice-president (see below).

Link: http://www.ercim.eu/about/members

Please contact:
Domenico Laforenza
IIT-CNR, ERCIM Vice-President
E-mail: domenico.laforenza@iit.cnr.it

ERCIM Symposium 2012

ERCIM is planning to establish a yearly scientific symposium held in conjunction with the ERCIM fall meetings with the goal to attract a larger audience to participate in ERCIM's activities. The first edition of the ERCIM Symposium will be held on 25 October 2012 as part of the ERCIM Fall Meetings in Sophia Antipolis, France, hosted by Inria.

The symposium will comprise scientific presentations as well as strategic panels taking a closer look at upcoming topics both on the scientific as well as policy level.

More information:
Stay informed about ERCIM activities through
http://www.ercim.eu, http://ercim-news.ercim.eu, @ercim_news
and the open ERCIM LinkedIn Group.

ERCIM Alain Bensoussan Fellowship Programme

The ERCIM Alain Bensoussan Fellowship Programme ERCIM offers fellowships for PhD holders from all over the world. The next deadline for application is 30 September. The Fellowship Programme is currently cofunded by The ERCIM "Alain Bensoussan" Fellowship Programme is co-funded by the European Commission FP7 Marie Curie Actions. More than 100 fellowships have already been granted under this COFUND scheme.

Who can apply?

The fellowships are available for PhD holders from all over the world.

What is the duration?

For the September 2012 deadline Fellowships are of 12 months duration spent in one institute.

Application deadlines:

Twice per year:
30 April and 30 September.

How to apply?

Only online applications are accepted. The application form will be online one month prior to the application deadline.

ERCIM Alain Bensoussan Fellowship Programme
MARIE CURIE ACTIONS
ERCIM
European Research Consortium for Informatics and Mathematics

ERCIM offers fellowships

- in Informatics and Applied Mathematics
- for PhD holders from all over the world
- in leading European research institutes

Application deadline twice per year
30 April and 30 September

fellowship.ercim.eu

Which topics/disciplines?

Topics cover most disciplines in computer science, information technology, and applied mathematics.

Where are the fellows hosted?

Fellows can be hosted at ERCIM member institutes only (the current ERCIM member institutes are listed on the back page of this issue). When an ERCIM member is a consortium (AARIT, CRCIM, PEG, PLERCIM, SARIT, SpaRCIM), the hosting institute might be any of the consortium's members. When an ERCIM Member is a funding organisation (FNR, FWO/FNRS), the hosting institute might be any of their affiliates.

What are the conditions?

- have obtained a PhD degree during the last eight years (prior to the application deadline) or be in the last year of the thesis work
- be fluent in English
- be discharged or get deferment from military service
- the fellowship is restricted to two terms (one reselection possible)
- have completed the PhD before starting the grant.
- a member institute cannot host a candidate of the same nationality
- a candidate cannot be hosted by a member institute, if by the start of the fellowship, he or she has already worked in this institute for a total of six months or more, during the last three years.

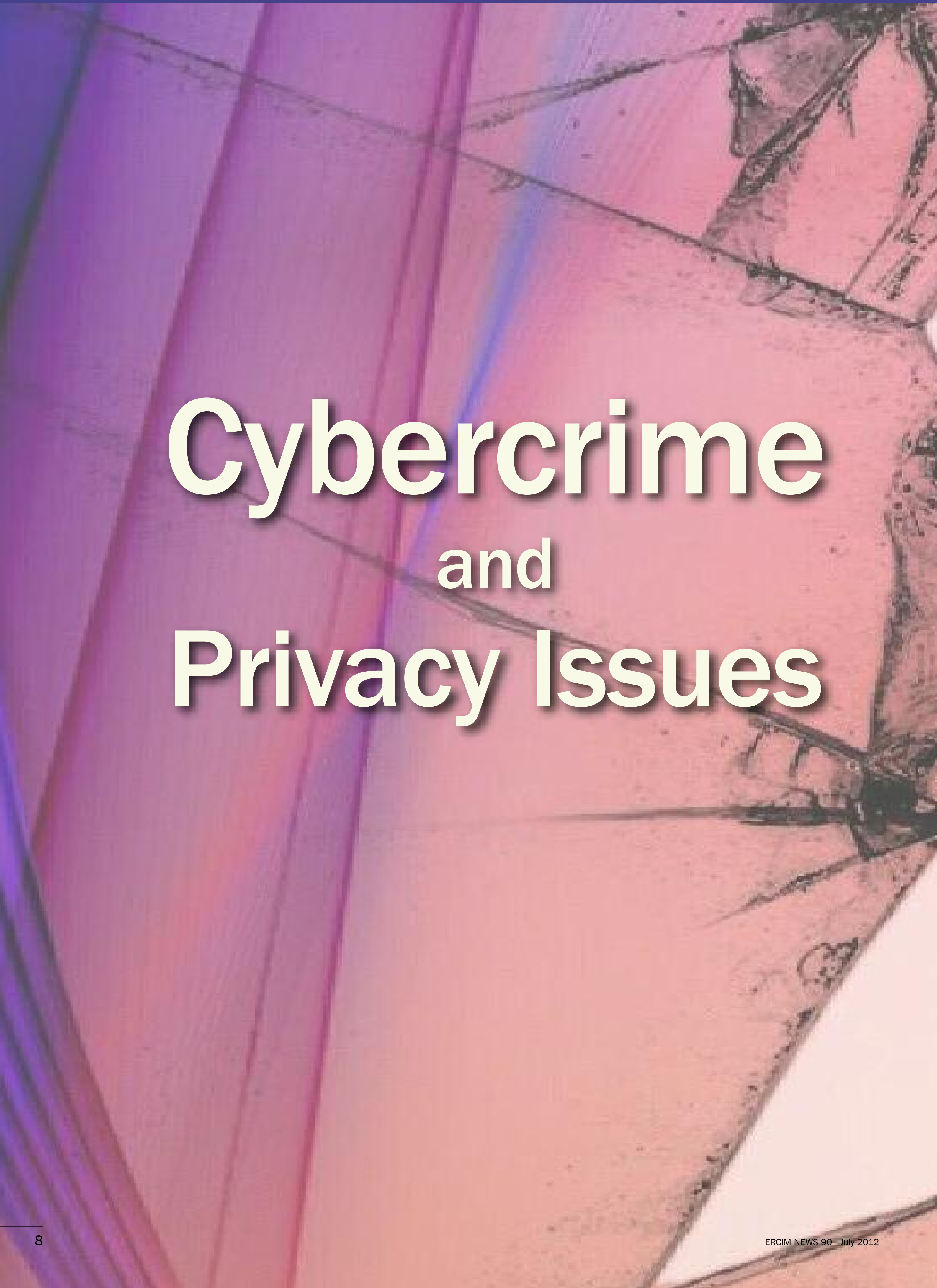
How are the fellows selected?

Each application is reviewed by scientists, and the criteria for selection are:

- scientific expertise of the applicant
- qualit y of scientific publications
- relevance of the fellow's research agenda
- interest/added-value for the ERCIM consortium
- previous mobility / professional experiences.

The number of available positions depends on the needs of the member institutes and their available funding.

More information: http://fellowship.ercim.eu/



Cybercrime and Privacy Issues

Introduction to the Special Theme

by Solange Ghernaouti-Hélie, Jens Tölle
and Jean-Jacques Quisquater

44 years ago Charles P. Lickson in a well-known paper "Privacy and the computer age" (IEEE Spectrum, October 1968, pp. 58-63) began his abstract with the prediction “By the year 2000, Americans could have computers and robots in the home - and virtually no privacy”. Now, in 2012, we could say better “virtually no privacy and a lot of cybercrimes”.

Cybercriminality has become a curse of society that affects everybody, nationally and internationally. Individuals, companies, institutions and governments may both become victims as well as (involuntary) helpers of cyber criminals. Inextricably associated with cyberspace, it is a reflection of the evolution of criminal practices that have adapted to the world of information and communication technologies.

Due to the world-wide distributed nature of today’s cyberspace, its infrastructure, services and user groups, criminals using this cyberspace for their activities form a severe challenge: This includes but is not limited to gathering of information on cybercrime related incidents, identification of proper persons in charge, or finding applicable laws. Often competence and responsibility are controversial.

The same holds for privacy: A multitude of cultures, different laws and different opinions makes it hard to agree on internationally standardised approaches.

This special edition of ERCIM represents a stage in the understanding of cybercriminality with reference to the need for the protection of digital privacy. It has to be recognised that the idea of digital privacy often suffers at the hands of information and communication technologies, and that personal data are intangible assets of great value, as much for legal entities as for criminals.

Although far from exhaustive and unable to cover all the aspects of both the fight against cybercriminality and the desire for the protection of privacy, this issue nonetheless presents an indication of various research projects and organisations that are tackling these problems and aims to inform readers about the kinds of technological measures being introduced to contribute towards better handling the vulnerabilities that can be exploited for malicious reasons. In fact cybersecurity is becoming a new and separate field of study that is ready for exploration in an interdisciplinary way, drawing upon the knowledge and techniques established in the fields of law, criminology, sociology, anthropology, economics, political science and digital technologies. This latter aspect is emphasised in this issue, although it should be stressed that the articles selected do not necessarily reflect the entirety of research activities across Europe and thus do not represent all of the academic institutions and research centres that are active and creative in this field.

The reader will find articles in this issue of ERCIM, covering different areas of research and showing the broad diversity of cybercrime and privacy. As far as research is concerned, the efficient understanding of cybercriminality needs now, more than ever, policies for supporting interdisciplinary research that encourages the decompartmentalization of traditional fields of research in favour of innovative projects in respect of the way of thinking about information security and about protecting assets. To this we should add a clear willingness to work together so that from an international perspective Europe will become a key player in the struggle against cybercriminality.

Invited article

The Cybercrime Ecosystem & Privacy Issues Main Challenges and Perspectives from a Societal Perspective

by Solange Ghernaouti-Hélie

An overview of the cybercriminal ecosystem

All the individuals and groups involved in cybercriminality, their ways of working, and the processes they have adopted to maximize their profits while minimising their risks of legal consequences; these elements go together to form an ecosystem. Like all ecosystems, this is lively, dynamic and undergoing permanent adaptation in order to exploit new opportunities in the marketplace, new vulnerabilities, new tools and new means of communication.

This ecosystem is a part of, and inseparable from, the ecosystem of the digital society. It possesses its own specific structures while involving legal users of the Internet and benefiting from the services that these provide. This is notably the case of entities that provide the facilities for financial transactions, such as, to name but two, Western Union or Liberty Reserve.

Cybercriminals are rational beings that follow the laws of the market and of supply and demand. They are above all criminals who have learned to extend their activities, knowledge and techniques into cyberspace. And in the same way as there exist a black market and a hidden economy in the physical world, the same can be found in cyberspace. These cybercriminal black markets work in the same manner as classical markets, with the objectives of performance and profitability, feeding the whole chain of cybercriminality and relying on the communications tools and opportunities for contacts provided by the Internet.

These markets use the same mechanisms, knowledge and tools as those activities linked to on-line advertising and legal e-commerce. They can be found at all stages of the performance of cybercrimes, of their preparation and their monetisation. In addition, the Internet contributes in a major way to realising their profits. Among the different possibilities offered by the black markets, it is possible to:

- Buy an on-line phishing kit, install it on a bulletproof server (classic hardware and software platforms), operate it (carry out phishing), collect the data gathered, and sell these through forums, on-line shops, and financial transaction services;
- Buy and sell exploits, malware and ransomware, software that allow cyberattacks to be carried out;
- Rent zombie machines and create and operate botnets;
- Buy and sell, wholesale or in small quantities, personal data such as banking details.

The stakes involved in protecting personal data and ensuring digital privacy

Cybercriminals know how to exploit personal data in order to optimise their activities and to reduce the risks of being held responsible for their own actions. Recent years have seen the development of a real economy based on the collection and sale of personal data, as well as the formation of a certain “criminal intelligence” around the use of these data. Without going into detail on these subjects, one can recognise the need for individuals and for society as a whole to have access to effective measures that will contribute towards protecting their personal data and their digital privacy, particularly with the objective of preventing, or at least limiting, the criminal use of these data.

At the same time we need to recognise that nowadays a lot of commercial organisations do use the personal data of Internet users within the framework of their entirely legal activities. This is true in general of the many service providers who propose services that are described as free. The Internet users pay in kind, indirectly, through supplying personal data, without necessarily having been aware of this or having given their express and informed permission. An important number of large Internet companies such as service and social networking platform providers take advantage of this situation to develop their economic models. They make large profits through commercialising and exploiting personal data, which users have either given freely or which has been collected without their knowledge.

To this kind of usage, which may be considered abusive by some, we can add the fact that these service providers that hold the personal data of their clients can themselves be the victims of cybercriminals (theft of data, infection and spread of malware, for example), and be an arena for cybercriminal activity insofar as their clients constitute numerous and attractive prey for the criminals.

In addition, all digital activities leave traces linked to personal data, which allows the permanent surveillance of Internet users by all kinds of operators.

This question should therefore not be seen solely in the perspective of the struggle against cybercriminality, but also in the perspectives of consumer protection (the consumers being Internet users) and of the protection of fundamental rights and of civil liberties, which include the freedom of speech, freedom of association, freedom of movement (the right to travel and to navigate freely on the Internet), the right to knowledge and information, and the right to respect for private life, family and correspondence. In order for these to be assured, it will be essential to be able to guarantee the protection of personal data and privacy, for these are elements that contribute to self-determination, to democracy, to liberty and, as a consequence, to human dignity. This all presupposes:

- Specific technological and judicial measures for protecting data;
- A genuine political and economic will in respect of the fair and honest

handling of personal data which will require the rethinking of economic models to ensure that personal data is not just considered as an asset to be traded;
- Coherent behaviour on the part of Internet users in respect of their data and of what they reveal about themselves on the Internet.

The place of the struggle against cybercriminality in the cybercriminal ecosystem

When considering the cybercriminal ecosystem, it is essential not to forget everyone else who is concerned by it, that is to say the individuals and the organisations who, depending on the circumstances, can find themselves the targets of, or the willing or unwilling participants in, cybercriminal acts. This latter distinction can be illustrated, for example, by the way that users can become a link in a criminal chain unwittingly as a result of fraud or manipulation. This is the case, for example, when a user's machine or an organisation's server acts as a relay or becomes a zombie member of a botnet used to carry out denial of service attacks on a third party. At the same time, a user can knowingly lend his machine to a botnet run by hacktivists, out of ideological, political, economic or religious convictions, for example. Public and private organisations, completely legally, can also be led to use the same weapons as cybercriminals in order to defend their interests. This can occur in the context of both offensive and defensive cybersecurity. An additional point to consider is that whenever an organisation represents certain values prized by the cybercriminals, as is the case of banks or commercial organisations offering on-line services, or whenever an organisation is responsible for the creation of assets, services, software or ICT or security solutions, that organisation by definition becomes a part of the cybercriminal ecosystem. Their presence in cyberspace, like that of Internet users who are very visible on social networks, for example, in some way explains the presence of cybercriminals and their activities.

The cybercriminal ecosystem would be incomplete if we did not include the police forces and judicial institutions that contribute in a very concrete operational way to combatting cybercriminality. They run criminal investigations and can be led to create honey pots. They use the same technical knowledge and the same tools as the cybercriminals. They can draw upon the specific technical knowledge of specially trained officers, of external civilian experts, or even of genuine cybercriminals, who may have repented or who simply have no other choice but to collaborate with the police. They can become full partners of the police, or act as informers, or actively work to deceive other cybercriminals, or track criminal activities and unmask their perpetrators, applying both their technical skills and their knowledge of the criminal environment.

As with classical investigations, this work requires a real police skill-set as it is not sufficient to be technically sound to be a good cybercrime investigator. They can sometimes have to operate undercover in order to infiltrate discussion forums on the black market, for example, or to infiltrate digital networks, which can sometimes be necessary in operations against Internet paedophiles.

The challenges for combatting cybercriminality

This would essentially consist of implementing technical, procedural, legal and organisational measures that would raise the number and quality of the difficulties in committing cybercrimes, increasing the level of risk for criminals and reducing the encouragements and the expected profits.

Such a programme would also include:
- The implementation of ICT infrastructures and services that are resilient and robust;
- The availability of comprehensive, transparent, manageable, effective, efficient security measures that are easy to implement, use and control;
- The global, integrated and effective strategic and operational management of information security as it concerns hardware, software, networks and cyberspace;
- The coherent and non-abusive use of information and communication technologies; and
- The faultless and ethical behaviour of all the members of the digital chain (users, managers, service providers).

There can be no fight against cybercriminality without a strong political and economic will to do so, without international agreements, without these agreements being respected, without the respect of fundamental human rights, without international cooperation and assistance, without considering the needs for justice, for peace, and for stability both in cyberspace and in the real world.

References:

- "Cybercrime, Cyberconflicts and Cybersecurity: a comprehensive approach", Ghernaouti-Hélie S, EPFL Press 2012
- "La cybercriminalité: le visible et l'invisible", Ghernaouti-Hélie S, Le Savoir Suisse 2009, ISBN 978-2-88074-848-7
- "In the world of Big Data, privacy invasion is the business model", http://news.cnet.com/8301-31322_3-57388097-256/in-the-world-of-big-data-privacy-invasion-is-the-business-model/, retrieved on 11 June 2012
- "A Global Treaty on Cybersecurity and Cybercrime", Schjolberg S and Ghernaouti-Hélie S, Second Edition, 2011. ISBN 978-82-997274-3-3

Please contact:

Solange Ghernaouti-Hélie
Director, Swiss Cybersecurity Advisory and Research Group
Faculty of Business and Economics
HEC, University of Lausanne
Lausanne, Switzerland
E-mail: sgh@unil.ch
http://www.hec.unil.ch/sgh

Measuring the Cost of Cybercrimes

by Michael Levi

Estimates of cybercrime costs are highly contested. We have become conditioned to believe that in order to generate control expenditure and powers to override privacy, very high attention-grabbing figures are needed. We were asked by the UK Ministry of Defence in 2011 to do a relatively 'quick and dirty' calculation to stimulate some serious analysis to counterbalance some of the high guesstimates currently in circulation, which have little general credibility. This attempt to dissect plausible data from scattered guesstimates was led by Ross Anderson from Cambridge and was co-authored by Chris Barton, Rainer Böhme, Richard Clayton, Michel van Eeten, Michael Levi, Tyler Moore, and Stefan Savage [1].

No study of the costs of cybercrime can be definitive. The spectrum is between a narrow summation of the known direct costs of detected crimes (perhaps even restricted to cases where a conviction has been obtained, because only then is criminality definitive), at one end, and speculative extrapolations from cases or sub-sets the dimensions of whose sets are unknown, at the other. In cyber, this is particularly complicated because it is a set of diverse acts representing mechanisms of crime commission, about which few organisations - whether victims or third parties like the police or vendors - compile data comprehensively or systematically. And unlike fraud, the costs of which one of us had reviewed previously [2], relatively little systematic effort had gone into measuring the costs of any sub-component of 'the cyber problem'. For each of the main categories of cybercrime we set out what is and is not known of the direct costs, indirect costs and defence costs – both to the UK and to the world as a whole, since the attribution of costs to particular countries is especially difficult in cyber. With global estimates, some fairly crude scaling based on GDP or in some cases, volumes of internet trade, have to be done to estimate costs to particular countries. Since the means (e. g., botnets) would not be around if there were not ends (e. g., phishing victims), we consider losses caused by the cybercriminal infrastructure as indirect by nature; irrespective of whether or not the legal framework formally criminalizes the means. We were more cautious than many others about the costs of IP espionage, since so little is known about both losses and whether external cyber-attacks or (as we suspect) internal corruption are the primary cause of those we do know about.

We distinguish carefully between traditional crimes that are now 'cyber' because they are conducted online (such as tax and welfare fraud); transitional crimes whose modus operandi has changed substantially as a result of the move online (such as credit card fraud); new crimes that owe their existence to the Internet; and what we might call platform crimes such as

Type of cybercrime	UK estimate	Global estimate	Reference period	Criminal revenue	Direct losses	Indirect losses	Defense cost
	in million US dollars						
Cost of genuine cybercrime							
Online banking fraud							
- phishing	16	**320**	2007	$\times^?$	$\times^?$		
- malware (consumer)	4	**70**	2010	$\times^{\downarrow}$	$\times^{\downarrow}$		
- malware (business)	6	**300**		$\times^{\downarrow}$	$\times^{\downarrow}$		
- bank technology countermeasures	50	**1 000**	2010				$\times^?$
Fake antivirus	5	**97**	2008-10	×	×		
Copyright-infringing software	**1**	**22**	2010	×			
Copyright-infringing music etc	7	**150**	2011	$\times^{\downarrow}$			
Patent infringing pharma	**14**	**288**	2010	×			
Stranded traveler scam	1	**10**	2011	$\times^{\downarrow}$			
Fake escrow scam	10	**200**	2011	$\times^{\downarrow}$			
Advance-fee fraud	**50**	1 000	2011	$\times^{\downarrow}$			
Cost of transitional cybercrime							
Online payment card fraud	**210**	4 200	2010				(×)
Offline payment card fraud							
- domestic	**106**	2.100	2010				$\times^{\downarrow}$
- international	**147**	2 940	2010				$\times^{\downarrow}$
- bank/merchant defense costs	120	**2 400**	2010				$\times^{\downarrow}$
Indirect cost of payment fraud							
- loss of confidence (consumes)	**700**	10 000	2010			$\times^?$	×
- loss of confidence (merchants)	**1 600**	20 000	2009			$\times^?$	×
PABX fraud	**185**	**4 960**	2011	×			$\times^{\downarrow}$
Cost of cybercriminal infrastructure							
Expenditure on antivirus	**170**	3400	2012			×	
Cost to industry of patching	50	**1000**	2010			$\times^?$	
ISP clean-up expenditures	**2**	40	2010		$\times^?$		
Cost to users of clean-up	**500**	10 000	2012		$\times^?$		
Defense costs of firms generally	500	**10 000**	2010			$\times^?$	
Expenditures on law enforcement	**15**	400	2010			×	
Cost of traditional crimes becoming 'cyber'							
Welfare fraud	**1 900**	20 000	2011	×	(×)		
Tax fraud	**12 000**	125 000	2011	$\times^?$	(×)		
Tax filing fraud		5 200	2010	×	(×)		

Table 1: Judgement on coverage of cost categories by known estimates.
Estimating costs and scaling: Figures in boldface are estimates based on data or assumption for the reference area. Unless both figures in a row are bold, the non-boldface figure has been scaled using the UK's share of world GDP unless otherwise stated in the main text. Extrapolations from UK numbers to the global scale should be interpreted with utmost caution. A threshold to enter this table is defined at $10m for the global estimate.
Legend: × : included, (×) : partly covered; with qualifiers ×↑ for likely over-estimated,
×↓ for likely underestimated, and ×? for high uncertainty.

the provision of botnets which facilitate other crimes rather than being used to extract money from victims directly.

As far as direct costs are concerned, we find that traditional offences such as tax and welfare fraud cost the typical citizen in the low hundreds of pounds/Euros/dollars a year; transitional frauds cost a few pounds/Euros/dollars; while the new computer crimes cost in the tens of pence/cents. In some cases, low production and distribution costs to criminals mean that direct social losses are roughly similar to criminal profits. For instance, UK consumers provided roughly $400,000 to the top counterfeit pharmaceutical programs in 2010 and perhaps as much as $1.2M per-month overall. UK-originated criminal revenue is no more than $14m a year, and global revenue, $288m. The five top software counterfeiting organisations have an annual turnover of around $22m worldwide. However, the indirect costs and defence costs are much higher for transitional and new crimes. For the former they may be roughly comparable to what the criminals earn, while for the latter they may be an order of magnitude more. As a striking example, the botnet behind a third of the spam sent in 2010 earned its owners around US$2.7m, while worldwide expenditures on spam prevention probably exceeded a billion dollars. Such defence expenditure is not necessarily irrational, but where crime is concentrated among a relatively small number of offenders, it makes sense to use criminal justice mechanisms to incapacitate the offenders. For example, the number of phishing websites, of distinct attackers and of different types of malware is persistently over-reported, leading some police forces to believe that the problem is too large and diffuse for them to tackle, when in fact a small number of gangs lie behind many incidents and a police response against them could be far more effective than telling the public to fit anti-phishing toolbars or to purchase antivirus software (though this might also be desirable). This is part of a much wider problem of attributing risks to patterns of offending. The legal-political problem is often how to take criminal justice action when suspects have been identified in a jurisdiction beyond ready reach! [3] Victimisation survey data suggest that cybercrime is now the typical volume property crime in the UK, and responses to it need to be mainstreamed. We do not claim that our analysis of the costs is more than a solid beginning in hotly disputed areas of which much is terra incognita. It is up to others to build upon these foundations: like the work of early cartographers, we may find that our map requires a lot more survey work.

Links/References:

[1] Measuring the Cost of Cybercrime
http://weis2012.econinfosec.org/papers/Anderson_WEIS2012.pdf

[2] The Nature, Extent and Economic Impact of Fraud in the UK. London: Association of Chief Police Officers. M. Levi, J. Burrows, M. Fleming, and M. Hopkins. (with the assistance of M. Matthews).
http://www.cardiff.ac.uk/socsi/resources/ACPO%20final%20nature%20extent%20and%20economic%20impact%20of%20fraud.pdf

[3] UK public and private sector expectations are explored further in M. Levi and M. Williams (forthcoming) eCrime Reduction Partnership Mapping Study, funded by Nominet Trust.

Please contact:

Michael Levi
Cardiff School of Social Sciences,
Wales, UK
E-mail: Levi@Cardiff.ac.uk

SysSec: Managing Threats and Vulnerabilities in the Future Internet

by Evangelos Markatos and Herbert Bos

For many years, cyber attackers have been one step ahead of the defenders. The asymmetric nature of the threat has led to a vicious cycle where attackers end up winning. SysSec, a new Network of Excellence in the area of Systems Security, attempts to break this vicious cycle and encourages researchers to work not on yesterday's attacks but on tomorrow's threats, to anticipate the attackers' next move and to make sure they are prepared.

Over the past decade we have seen a large number of cyber attacks on the Internet. Motivated by financial profits or political purposes, cyber attackers usually launch attacks that stay below the radar, are difficult to detect, and exploit the weakest link: the user. We believe that the core of the problem lies in the nature of cyber security itself: in the current practice of cyber security, most defenses are reactive while attackers are by definition proactive. Cyber security researchers usually chase the attackers trying to find one more defense mechanism for every newly created attack. Thus, we are facing an asymmetrical threat: while attackers have all the time in the world to choose when and where to strike minimizing their cost, defenders must respond fast, within narrow time constraints, and at a very high cost. Each new round of attack-and-defense drains energy from the defenders, leading them down a vicious cycle which will eventually wear them out. It seems that the only way to build effective defenses is to break this cycle, by changing the rules of the game, by anticipating the moves of the attackers, and by being one step ahead of them, through (i) identifying emerging vulnerabilities, and (ii) working towards responding to possible attacks before they appear in the wild. In this aspect, the recently created SysSec Network of Excellence takes a game-changing approach to cyber security: instead of chasing the attackers after an attack has taken place, SysSec studies emerging

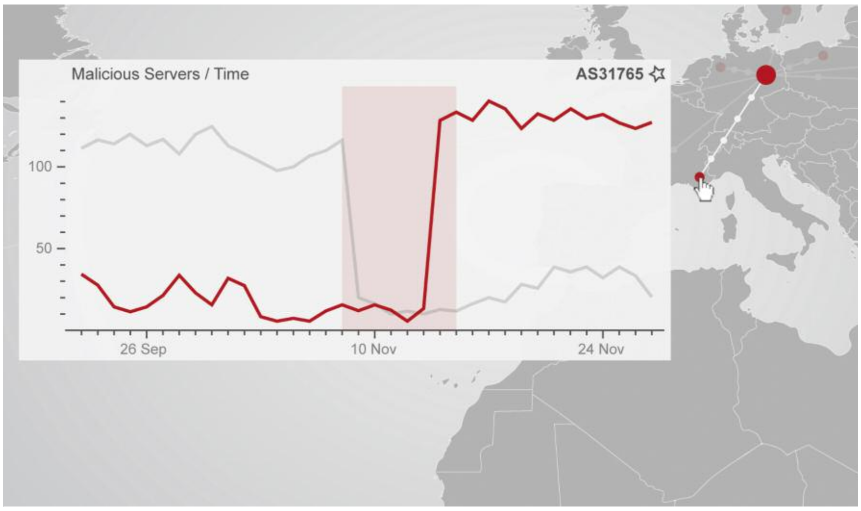


Figure 1: SysSec's BURN interface visualises malicious activities in autonomous systems---in this case, the number of malicious servers as a function of time for a network in Germany exhibits a sudden drop, whereas we find a specular sudden step in a network in France. BURN makes it easy to correlate this type of events visually.

threats and vulnerabilities ahead of time. The network's main thrusts are to identify a roadmap to work on threats and to build infrastructure to boost education in system security—to provide the expertise needed to deal with these emerging threats.

Roadmap

With the collaboration of the research community, SysSec has already produced a research roadmap (http://syssec-project.eu/roadmap1) which outlines some of the important areas the community feels we should focus on. In the first year, the project selected five categories:

1. Privacy. SysSec urges researchers to investigate how to protect users against sophisticated attacks that aim to disclose their personal information. For example, it is important to promptly detect functionalities that can be abused to correlate data available in public records and de-anonymize user accounts in many online services.
2. Targeted attacks. It is important for researchers to develop new techniques to collect and analyze data associated with targeted attacks. The lack of available datasets, in addition to the limitation of the traditional analysis and protection techniques, is one of the current weak points of the war against malware. The problem is often to find the needle of the targeted attack in the haystack of the traditional attacks perpetuated every day on the Internet. In addition, researchers should focus on new defense approaches that take into account alternative factors (such as monetization), and large scale prevention and mitigation (e.g., at the Internet Service Providers (ISP) level).
3. Security of emerging technologies, in particular the cloud, online social networks, and devices adopted in critical infrastructures (like smart meters). Security in new and emerging technologies before it is too late is one of the main priorities of the system security area. In this direction, it is important to sponsor activities and collaboration between academia and the industrial vendors to maximize the impact of the research and reduce the time required for the analysis and the experiments.
4. Mobility: develop new tools and techniques that can be deployed in current smartphone systems to detect and prevent attacks against the device and its applications.
5. Usable security: We believe that a study of the usability of security measures is important and it will become even more critical in the future. If we want to progress in this direction, we need interdisciplinary efforts that bring together experts from different fields (including engineering, system security, psychology, etc.).

With the help of experts organized in working groups, SysSec updates its roadmap yearly to reflect new threats and priorities.

Education

Having realized the lack of educational material in the area, SysSec further aims to establish a center for academic excellence in the area and has started designing a common curriculum on cyber security, focusing mostly on the production of slides and lab exercises, which are particularly hard to design and set up. A first version of the curriculum along with course material is expected to be ready by September 2012. It will be open to universities throughout Europe and will help to set up a state of the art cyber security curriculum to train the next generation of experts.

We underline that besides SysSec several other projects aim to map the research landscape in cyber security. However, with a clear focus on system security and the development of usable course material, we believe SysSec occupies a unique and valuable niche. SysSec may be contacted at contact@syssec-project.eu, may be followed in twitter (twitter: syssecproject) and may be found in Facebook (http://www.facebook.com/SysSec).

References:

Privacy-Preserving Social Plugins

[1] G. Kontaxis, M. Polychronakis, A. D. Keromytis and E; P. Markatos. "Privacy-Preserving Social Plugins", In the Proceedings of the 21st USENIX Security Symposium, 2012.

[2] F. Maggi, A.Volpatto, S. Gasparini, G. Boracchi, S. Zanero. "POSTER: Fast, Automatic iPhone Shoulder Surfing". In the Proceedings of the 18th ACM/SIGSAC Conference on Computer and Communications Security (CCS), 2012.

[3] C. Rossow, C. J. Dietrich, C. Kreibich, C. Grier, V. Paxson, N. Pohlmann, H. Bos and M. van Steen. "Prudent Practices for Designing Malware Experiments: Status Quo and Outlook". In the Proceedings of the 33rd IEEE Symposium on Security & Privacy (Oakland), 2012.

Please contact:

Herbert Bos, VU University
Amsterdam, The Netherlands
Tel: +31-20 598 7746
E-mail: HerbertB@cs.vu.nl

Evangelos Markatos
FORTH-ICS, Greece
Tel: +30 2810391655
E-mail: contact@syssec-project.eu

Understanding the Role of Malware in Cybercrime

by Juan Caballero

At the core of most cybercrime operations is the attacker's ability to install malware on Internet-connected computers without the owner's informed consent. The goal of the MALICIA project is to study the crucial role of malware in cybercrime and the rise in recent years of an "underground economy" associated with malware and the subversion of Internet-connected computers.

Cybercrime, criminal activity conducted via computers connected to the Internet, is a growing threat for developed regions like Europe where nearly three quarters of households and a large number of the infrastructures are connected to the Internet, and an increasingly number of services and transactions happen online.

At the core of most cybercrime operations is the attacker's ability to install malicious programs (ie malware) on Internet-connected computers without the owner's informed consent. Malware includes bots, viruses, trojans, rootkits, fake software, and spyware. Malware enables attackers to establish a permanent presence in the compromised computer and to leverage it for their cybercrime operations. The target of these operations may be the compromised computers themselves eg stealing an organization's intellectual property or a user's banking credentials, or third parties. In the latter case, the compromised computers are simply assets, which the attacker employs to launch malicious activities such as sending spam, launching denial-of-service (DoS) attacks, faking user clicks on online advertisements (ie click-fraud), or simply as a stepping stone to hide its location.

The goal of the MALICIA project at the IMDEA Software Institute is to study the crucial role of malware in cybercrime and the recent rise of a far-reaching "underground economy" associated with malware and the subversion of Internet-connected computers. Gone are the days where attackers compromised computers and built malware to show off their skills to peers. These days, the malware ecosystem revolves around cybercrime and the monetization of compromised computers.

As the malware ecosystem has grown larger and more profitable, specialization has come to the marketplace. Attackers have understood that tackling the entire value-chain from malware creation to monetization poses a daunting task requiring highly developed skills and resources. As a result, specialized services have been created at all stages in the malware-monetization chain, such as toolkits to automate the construction of malware, program encryption tools to evade antivirus (AV) software, "bullet-proof" hosting, and forums for buying and selling ill-gotten gains. Specialized services lower the barrier to entering the malware ecosystem. However, defenders can also take advantage of specialization since disrupting the specialized services disrupts the different malware operations using them.

As a first step in the MALICIA project, we have collaborated with researchers at the University of California, Berkeley and the International Computer Science Institute to investigate the commoditization of malware distribution in the form of pay-per-install (PPI) services. PPI services offer criminals a simple way to outsource the distribution of their malware. The clients provide their malware to the PPI service and select the number of desired installations (called installs) in each geographical area. The PPI service takes care of installing the malware on compromised computers in exchange for a small fee that ranges from $180 for a thousand computers in some European countries and the US, down to $7 for a thousand computers in Asia.

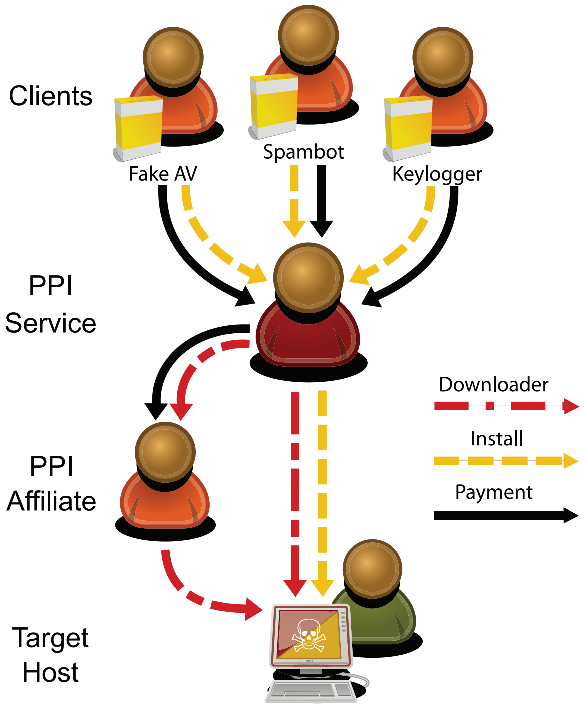


Figure 1: The Pay-Per-Install market

To satisfy the clients' demand for installs, the PPI provider typically outsources malware distribution to third parties called affiliates. PPI providers pay affiliates for each compromised computer, acting as a middle man that sells installs to the clients while buying installs from affiliates. Each affiliate may specialize in some specific malware distribution method (eg bundling malware with a benign program and distributing the bundle via file-sharing networks; exploiting web browsers through drive-by-downloads; or social engineering). The PPI service gives each affiliate a downloader program customized with a unique affiliate identifier. When the affiliate installs the downloader in a compromised computer, the downloader connects back to the PPI service to download the client programs. After installing the client programs on the compromised host, the downloader reports the affiliate identifier and the affiliate is credited with an install.

To understand the PPI market we infiltrated four PPI services. For this, we developed infrastructure enabling us to (1) interact with PPI services by mimicking the protocol interactions they expect to receive from affiliates, and (2) collect and classify the malware being distributed by the PPI services. Using this infrastructure we harvested over a million malware programs and classified them by malware family as well as monetization methods. Our analysis revealed that of the world's top 20 malware families, 12 employed PPI services for their distribution. It also revealed that some malware families exclusively target the US and a variety of European countries. The monetization methods in use are varied including: spam, installing fake antivirus software, information-stealing, denial-of-service, click-fraud, and adware.

Much remains to be learnt about the malware ecosystem and the specialized economy supporting cybercrime. Our current work strives on deepening our understanding of other parts of the ecosystem. One overarching goal is evolving malware analysis from understanding what a malware program does, to also cover why it does it, ie what role the malware program plays in the cybercrime operation where it is used.

References/Link:

- "Measuring Pay-per-Install: The Commoditization of Malware Distribution", J.Caballero, C. Grier, C.Kreibich, and V. Paxson. In Proc. of the 20th USENIX Security Symposium, San Francisco, CA, August 2011.
- "Most Malware Tied to 'Pay-Per-Install' Market", B. Krebs, MIT Technology Review, Thursday, June 9, 2011. http://www.technologyreview.com/news/424241/most-malware-tied-to-pay-per-install-market/

Please contact:
Juan Caballero, IMDEA Software Institute, Spain
E-mail: juan.caballero@imdea.org

Peering into the Muddy Waters of Pastebin

by Srdjan Matic, Aristide Fattori, Danilo Bruschi and Lorenzo Cavallaro

Advances in technology and a steady orientation of services toward the cloud are becoming increasingly popular with legitimate users and cybercriminals. How frequently is sensitive information leaked to the public? And how easy it is to identify it amongst the tangled maze of legitimate posts that are published daily? This underground bazaar is, after all, under the eyes of everyone. Do we have to worry about it and can we do anything to stop it?

Pastebin applications, also known simply as “pastebin”, are the most well-known information-sharing web applications on the Internet. Pastebin applications enable users to share information with others by creating a paste. Users only need to submit the information to be shared and the service provides an URL to retrieve it. In addition to being useful for sharing long messages in accordance with policies (e.g., Twitter) and netiquette (IRC chats), one of the main features that make pastebin appealing is the possibility of anonymously sharing information with a potentially large crowd.

Unfortunately, as along with the legitimate use of such services comes their inevitable exploitation for illegal activities. The first outbreak occurred in late 2009, when roughly 20,000 compromised Hotmail accounts were disclosed in a public post. Many other sensitive leaks followed shortly thereafter, but it is with the illegal activities of the hacker groups Anonymous and LulzSec that such security concerns reached a much wider audience [1].

To shed interesting insights on the underground economy, we, Royal Holloway, University of London and University of Milan, jointly developed a framework to automatically monitor text-based content-sharing pastebin-like applications to harvest and categorize (using pattern matching and machine learning) leaked sensitive information.

We monitored pastebin.com from late 2011 to early 2012, periodically downloading public pastes and following links to user-defined posts. We recorded a diverse range of categories of sensitive or malicious information leaked daily: lists of compromised accounts, database dumps, list of compromised hosts (with backdoor accesses), stealer malware dumps, and lists of premium accounts.

The list of compromised accounts (i.e., username and password pairs) is the most commonly recorded stolen sensitive information (685 posts with 197,022 unique accounts). Such lists are often packed with references to where these accounts were stolen and the websites where they would be valid, giving miscreants (or just random curious readers) an easy shot. Such information enables us to shed some light on previous security trends and weaknesses [2] (e.g, password strengths and credential reuse). For instance, more than 75% of such passwords were cracked in a negligible amount of time, pointing out that users still rely on poorly chosen or weak passwords.

Similarly, posts of leaked database dumps often include references to the attacked servers, precise information on the exploited vulnerability and clear indications of the tools used to perform the attack, providing interesting insights into the attackers’ methods.

Posts containing leaked information about compromised servers (104 posts

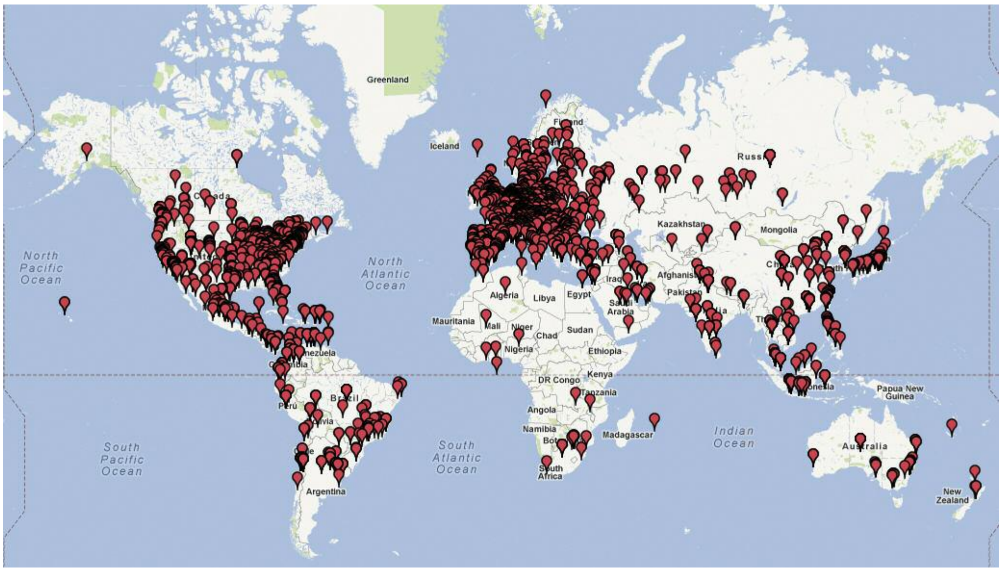


Figure 1: Geographical distribution of shells

with 5,011 unique accounts) include lists of URLs with recurring patterns (e.g., webdav, shell, dos). Our analysis shows that such PHP-written shells are generally aimed at performing UDP-based DoS attacks.

Information leaked by malware was responsible for 121 posts with 12,036 unique accounts. Such posts report very precise information associated with the leaked credentials, i.e., the URL of the website for which the account is valid, the program from which they were stolen, an IP addresses, a computer name and a date.

Finally, posts of leaked premium website accounts contain lists of username and password used to access web applications that provide enhanced features for paying customers (892 posts with 239,976 unique accounts). Unsurprisingly, the two commonest categories of premium accounts refer to pornography and file sharing websites.

As previous researchers have done [2], we evaluated the potential value of this sensitive information on the black market [3]; prices and values are reported in Table 1.

As outlined above, some leaked posts linked to shell installed on compromised servers. To better understand the threat posed by the public disclosure of such information, we evaluated the bandwidth capacity (using a geo-location database) these shells may generate in a DDoS attack. Out of more than a hundred shell-related posts, we extracted roughly 31,000 shell-related URLs (5,011 unique, 4,784 of which valid). Such shells are installed on servers located in 118 different Countries (as shown in Figure 1), with the top five referring to USA (1074), Germany (629), The Netherlands (219), France (166), and UK (164). The aggregate computed bandwidth is 23.3Gbps, comparable to that of a small botnet.

Our analysis reported 121 posts containing stealer malware dumps. We identified roughly 14,000 dumps (12,036 of which were unique). Owing to the structured nature of these dumps, it was possible to gather precise statistics (omitted from this article due to space constraints). Most of the websites were about gaming, social networking, and file sharing.

In conclusion, our ongoing research effort showed that sensitive information is easily and publicly leaked on the Internet. The automatic identification of such information is not only an interesting research topic, as it sheds insights on underground economy trends, but, if properly enforced, it may allow us to detect and contain the damage caused by malicious leaks.

Link:

[1] LulzSec, "Fox.com hack", http://pastebin.com/Q2xTKU2s, 2011

References:

[2] Brett Stone-Gross, Marco Cova, Lorenzo Cavallaro, Bob Gilbert, Martin Szydlowski, Richard Kemmerer, Christopher Kruegel, and Giovanni Vigna, Giovanni, "Your botnet is my botnet: analysis of a botnet takeover", Proceedings of the 16th ACM conference on Computer and Communications Security, 2009

[3] Symantec Corporation, "Symantec Internet Security Threat Report 2010", 2010

Please contact:

Lorenzo Cavallaro
Royal Holloway, University of London
Tel: +44 1784 414381
E-mail: lorenzo.cavallaro@rhul.ac.uk

Item	Price Range	Quantity	Value
Email Accounts	\$1—\$18	102,522	\$78,808—\$1,418,558
Email Addresses	\$1/MB—\$20/MB	1,851,552	\$1.8—\$36
Attack Tools	\$5—\$650	~5,000	~\$25,000
Premium Accounts	\$10—\$110	239,976	\$2,399,760—\$26,397,360

Table 1: Prices and values of goods on the black market

User Data on Androïd Smartphone Must be Protected

by Radoniaina Andriatsimandefitra, Valérie Viet Triem Tong, and Ludovic Mé

"In the world of mobile, there is no anonymity," says Michael Becker of the Mobile Marketing Association, an industry trade group. In recent work, Enck and colleagues have used information flow monitoring on a mobile device to show that, on average, over two thirds of the most popular applications of an Android market were responsible for data leakage [1]. We believe data leakages are mainly due to the intrinsic limitations of Android's security mechanisms. Here we describe "Blare", a tool that detects Androïd data leakages.

Android is an operating system for mobile devices. Because of its quick and wide-scale adoption, it has become the target of malicious applications which continue to increase in number. This increase is alarming given that more and more people are relying on such devices both for personal and professional use. A protection system is essential but unfortunately, existing mechanisms fail to efficiently protect sensitive data located on a smartphone. We argue that leakages are mainly due to intrinsic limitations to Android's security mechanisms that rely heavily on access control systems and, as such, offer no possible control over access to a piece of data once it has left its original location.

In CIDRe, a joint project team with SUPELEC and INRIA, we have designed and developed "Blare", a Linux Intrusion Detection System (IDS). Blare makes use of tainting to monitor occurring information flows and detect illegal ones in the context of a predefined security policy [2]. Blare relies on the LSM framework, a patch for the Linux kernel that inserts "hooks" at every point in the kernel where a user-level system call generates an information flow. Well-known security modules such as SELinux, AppArmor, Smack and TOMOYO also rely on LSM. Blare maintains two security tags for each object of the operating system (files, processes, etc.). The first tag lists the sensitive data that were used to produce the current content of the object. The second tag details which data mixture is allowed to flow in the object (ie describes the security policy that applied to the object). The legality of an information flow is thus easily established by comparing the tags' values. When the values do not match, Blare raises an alert.

Having an efficient linux implementation of our tool, we then studied its applicability in the Androïd context. We have proposed an entire information flow policy dedicated to the protection of data usually located on a smartphone (eg contact list, geolocalization), and implemented an Androïd version of Blare. Our measurements on application overhead have offered encouraging results and we are able to detect violation of integrity or confidentiality of data.

To test our ability to detect attacks against integrity, we have used BaseBridge, an Android malware whose purpose is to install other malicious applications on the phone, thus violating the integrity of the system. To detect these malicious installations, we tagged each piece of data coming from the analysed sample and monitored its propagation within the system. As expected, alerts were raised by Blare. Using the alerts, we then built a graph that described the propagation of tainted data within the system, taking into account the timestamp of each alert. The graph showed that the application meant to be installed came from a file owned by the basebridge sample and that its content was accessed/stored by different containers in a way that indicates its installation and execution.

To test our ability to detect attacks against confidentiality, we exploited two well-known vulnerabilities. The first is related to the Android browser and allows data leakage out of the device. The second is related to the Androïd Skype app that leaks user-related data. We exploited these two vulnerabilities to leak Skype data through the browser to a remote entity. We tagged sensitive data container prior to the attack. Once our attack launched, Blare raised meaningful alerts. The alerts clearly showed that the default browser had read files with sensitive content and also leaked this sensitive data to the remote entity. Using the alerts, we built a graph that describes how sensitive tagged-data were leaked. The graph indicated that the browser first accessed the sensitive tagged-data stored in Skype directory and then wrote this data inside a socket to send it to the remote entity.

To conclude, Blare has proved efficient in detecting attacks, and especially data leakage, in the Android context.

Link:
http://www.blare-ids.org

References:

[1] W. Enck, P. Gilbert, B. Gon Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth, "Taintdroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones" in Proc. of the USENIX Symposium on Operating Systems Design and Implementation (OSDI), 2010.

[2] V. Viet Triem Tong, A. Clark, and L. Mé, "Specifying and Enforcing a Fine-Grained Information Flow Policy: Model and Experiments" in Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications, 2010.

[3] R. Andriatsimandefitra, S. Geller, and V. Viet Triem Tong, "Designing Information Flow Policies for Android's Operating System" in Proceedings of the IEEE Internation Conference on Computer Communications (ICC), 2012.

Please contact:

Ludovic Mé
EPC SUPELEC/Inria CIDRE, Rennes, France
Tel: +33 299844500
E-mail: ludovic.me@supelec.fr

i-Code: Real-Time Malicious Code Identification

by Stefano Zanero, Sotiris Ioannidis and Evangelos Markatos

A European Union project aims to provide a much needed toolkit for forensic analysis of network-spreading malicious code.

Networks are currently infested with malicious programs propagating from one computer to another. Referred to by colourful names such as worms, viruses, and shellcodes, these programs seek to penetrate and compromise remote computers by exploiting their vulnerabilities. Once a remote computer is compromised, it can be used for a wide variety of illegal activities including blackmailing, Denial of Service attacks, illegal material hosting, sending of SPAM and fraud.

i-Code is a two-year-long research project aimed at realizing an integrated real-time detection and identification toolbox for malicious code. The toolbox (complete with an integrated console) can help network administrators and forensic analysts wishing to investigate an incident involving malicious code.

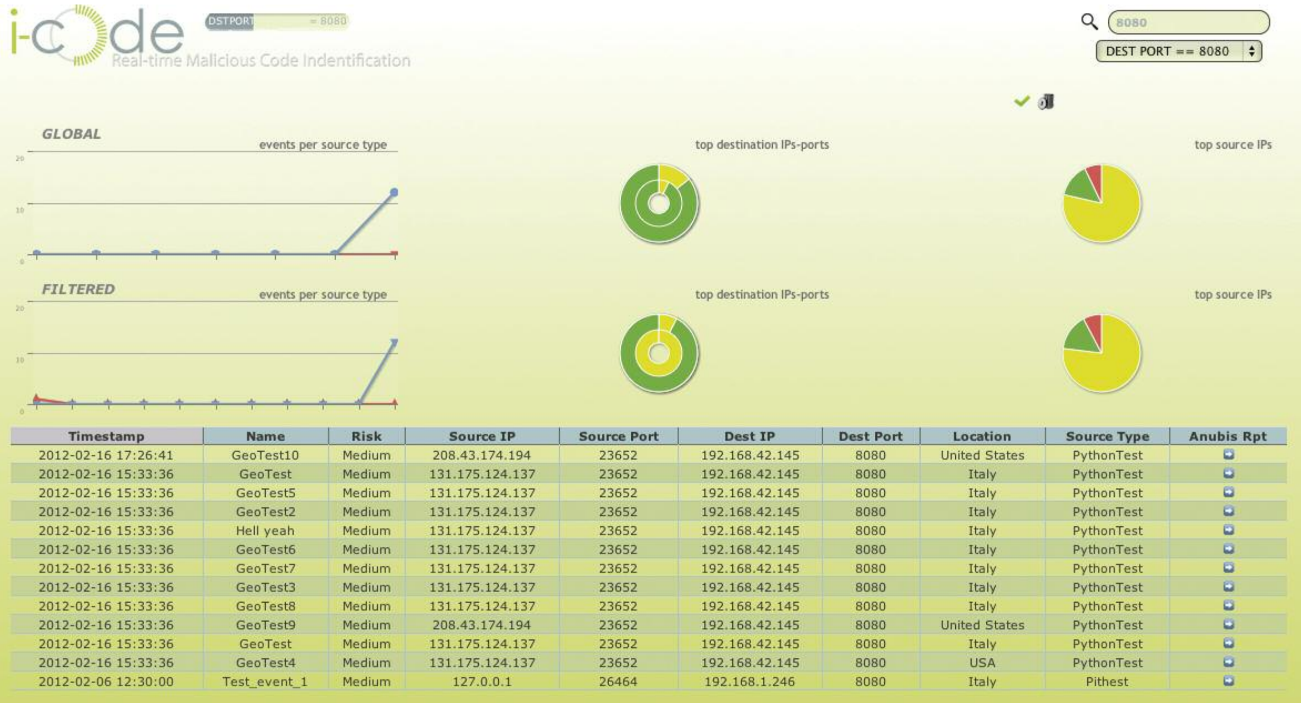
Having identified the challenge that the increasing network speeds pose to attack detection, i-Code addresses this challenge by developing an I/O architecture which avoids common bottlenecks by reconfiguring datapath logic at application load time to match workload and exploit special-purpose hardware. The two components of the I/O architecture that are crucial for performance are processing and buffering. For processing, i-Code reuses the well-known streams and filters model and for buffering, i-Code employs a buffer management system where all live data are kept in coarse-grain ring buffers, and buffers are shared long-term between protection domains.

The detection approach of i-Code is three-pronged, with two network-level detectors (NEMU and Argos) and a host level detector (AccessMiner), integrated by a forensic console which also correlates their results and glues them together with shellcode analysis provided by the Anubis sandbox.

NEMU [1] is a tool that performs network-level emulation, a heuristic detection method that scans network traffic to detect polymorphic attacks. NEMU uses a CPU emulator to dynamically analyse every potential instruction sequence in the inspected traffic, and attempts to identify the execution behaviour of certain malicious code classes, such as self-decrypting polymorphic shellcode. Attackers trying to hide their malware inside ordinary-looking incoming network packets, are easily discovered by NEMU. Complemented by Anubis, a dynamic malware analysis sandbox, NEMU is able not only to detect, but also accurately classify incoming attacks.

Argos [2] is a full and secure system emulator designed for use in honeypots. It is based on Qemu but has been extended to detect remote attempts to compromise the emulated guest operating system. Using dynamic taint analysis, it tracks network data throughout execution and detects any attempt to use them in an illegal way. When an attack is detected the memory footprint of the attack is logged.

AccessMiner [3] is a tool developed to analyse system calls collected on hosts that run applications for regular users on actual inputs, and differentiate them from malware system calls. It has been designed for large scale collection and centralized analysis on real world networks.



Timestamp	Name	Risk	Source IP	Source Port	Dest IP	Dest Port	Location	Source Type	Anubis Rpt
2012-02-16 17:26:41	GeoTest10	Medium	208.43.174.194	23652	192.168.42.145	8080	United States	PythonTest	➡
2012-02-16 15:33:36	GeoTest	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest5	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest2	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	Hell yeah	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest6	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest7	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest3	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest8	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest9	Medium	208.43.174.194	23652	192.168.42.145	8080	United States	PythonTest	➡
2012-02-16 15:33:36	GeoTest	Medium	131.175.124.137	23652	192.168.42.145	8080	Italy	PythonTest	➡
2012-02-16 15:33:36	GeoTest4	Medium	131.175.124.137	23652	192.168.42.145	8080	USA	PythonTest	➡
2012-02-06 12:30:00	Test_event_1	Medium	127.0.0.1	26464	192.168.1.246	8080	Italy	Pithest	➡

Figure 1: A screenshot of the i-code console. The events shown are fictional and the IP addresses not real.

Anubis is a dynamic malware analysis system based on an instrumented Qemu emulator. It is offered as an open service through a public website, where users can submit binaries for analysis, and receive a report that describes the behaviour of the sample in a human-readable way. For i-Code Anubis was extended to support the analysis and classification of shellcode.

The console is designed to collect events generated by these systems, pass the resulting shellcode on to the Anubis sandbox for analysis, and integrate the results in an easy-to-use view. It is also designed to be easily extensible with further detection systems through the use of open communication standards.

The final results of the project were presented in Brussels in June 2012 in a conference, attended by over 40 members of the European forensics community. Besides project results, several open source or research tools for network monitoring and incident analysis were presented.

Links:
i-Code: http://www.icode-project.eu
Anubis: http://anubis.iseclab.org
Anubis Shellcode Analyzer:
http://shellcode.iseclab.org/
NEMU:
http://www.ics.forth.gr/dcs/Activities/papers/nemu.wdfia09.pdf
Argos: http://www.few.vu.nl/argos/

References:

[1] M. Pol ychronakis, E. P. Markatos, and K. G. Anagnostakis. Emulation-based detection of non-self-contained polymorphic shellcode. In Proceedings of the 10th International Symposium on Recent Advances in Intrusion Detection (RAID), September 2007.

[2] G. Portokalidis, A. Slowinska, and H. Bos. 2006. “Argos: an emulator for fingerprinting zero-day attacks for advertised honeypots with automatic signature generation”. In Proc. of the 1st ACM SIGOPS/EuroSys European Conf. on Computer Systems 2006 (EuroSys '06). ACM, New York, NY, USA, 15-27.

[3] A. Lanzi et al 2010. “AccessMiner: using system-centric models for malware protection”. In Proc. of the 17th ACM Conf. on Computer and communications security (CCS '10). ACM, New York, NY, USA, 399-412.

Please contact:

Stefano Zanero
Politecnico di Milano, Italy
Evangelos Markatos
ICS-FORTH, Greece
E-mail: contact@icode-project.eu

A Scalable Approach for a Distributed Network of Attack Sensors

by Jan Gassen and Elmar Gerhards-Padilla

Today’s computer systems face a vast array of severe threats that are posed by automated attacks performed by malicious software as well as manual attacks by individual humans. These attacks not only differ in their technical implementation but may also be location-dependent. Consequentially, it is necessary to join the information from heterogeneous and distributed attack sensors in order to acquire comprehensive information on current ongoing cyber attacks.

The arms race between cyber attackers and countering organizations has spawned various tools that are even capable of detecting previously unknown attacks and analyzing them in detail. Owing to the heterogeneity of possible cyber attacks, various tools have been developed to deal with certain classes of attacks, for instance, “server honeypots” masquerade as regular production systems waiting to be probed and attacked directly over the network. These systems are therefore able to detect network based attacks against vulnerable services but cannot detect attacks that are performed against client applications, for example by using malicious documents. Moreover, even within one class of attack, different detection tools may be needed since existing tools are generally only able to detect a subset of possible attacks that may occur within a particular class. Finally, the information gathered by individual tools can be further enhanced by a set of passive tools generating additional details about the attacker’s origin or to examine the attacker’s operating system by fingerprinting the observed network connection.

In order to allow a comprehensive detection of cyber attacks, the Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE in conjunction with the University of Bonn is developing a distributed architecture for heterogeneous attack sensors in different geographic locations and networks. The major goal of the project is to create a scalable architecture with large amounts of sensor locations that can be easily extended by additional as well as new types of sensors.

Creating a scalable architecture primarily requires distributing the network load as well as the required computational effort to different resources. Therefore, the entire set of applied sensors is broken down into different independent sites, whereas every site contains a particular set of sensors (Figure 1). Each site is responsible for generating extensive information about attacks observed at the according network or geographical location. The core component of a single site is a central correlation server that receives information from the different independent sensors and automatically aligns information that belongs to the same incident. During this process, redundant or duplicate information is deleted before the resulting data is streamed to connected clients in unified JSON format. These clients are used to further process the collected information and can be

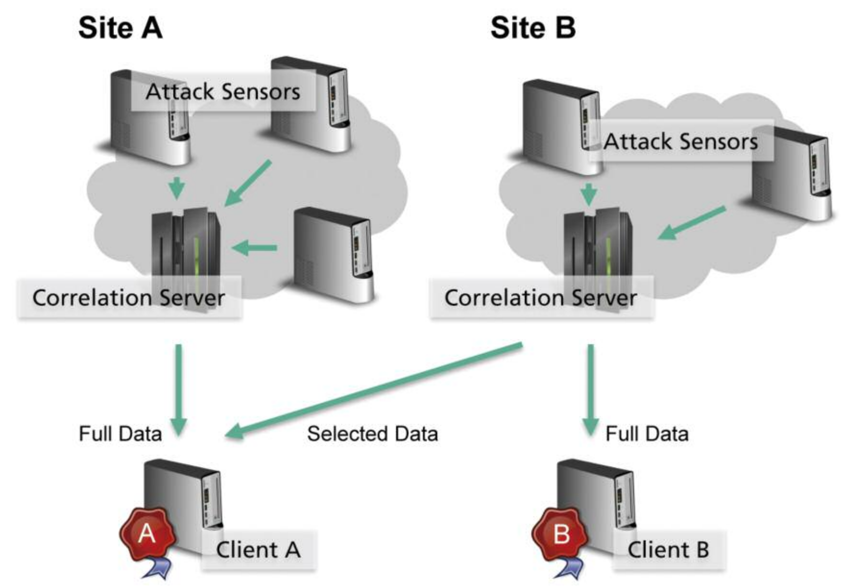


Figure 1: Overview of the distributed architecture for distributed and heterogeneous attack sensors.

located independently from the connected sites.

As a result of the preprocessing within the individual sites, a connected client only receives relevant and pre-parsed data that can be directly used for further processing. To store the received data stream, the client uses the distributed and schema-less database MongoDB. This allows deployment of the database on multiple resources to achieve the performance as needed for individual purposes. Moreover, the use of JSON data and a schema-less database design allows almost arbitrary data to be stored without the need to adjust the database, the client or the correlation server. Therefore, new sensors can be added to existing sites and new sites with different sensors can be added to a certain client without any adjustments.

The more sites that are deployed within different networks or geographic locations, the more comprehensive the information gathered about ongoing attacks. It is therefore desirable for organizations to share data with each other, thus reducing the need for their own resources and according maintenance costs. On the other hand, involved organizations may only be willing to share a particular subset of their registered data. To this end, the correlation server within each site allows certificate-based authentication, whereby the certificates are also used to specify the information to be sent to the according client.

To be able to manage and maintain administrated sensors within one or more sites from a single point, a controlling application has been developed that can be deployed on every resource involved. On each resource, the controlling application observes running processes and is able to perform simple countermeasures in case of detected errors. If the controlling application is not able to resolve the observed issue automatically, the administrator is notified automatically to take further action. Furthermore, all controlling applications participate in a peer-to-peer network by using JGroups, allowing the administrator to join the network and issue commands to all participating resources.

A remaining challenge in observing different attacks from different locations is the increasing amount of data. While the data can be handled automatically by distributed databases, it remains hard for human analysts to extract particularly relevant data. Therefore, future work is focused on automatic clustering of incoming data in order to group similar attacks. Thus, newly registered attacks can be directly compared with similar known attacks, simplifying their identification and reducing the workload of human analysts.

Link:
http://www.fkie.fraunhofer.de/en/research-areas/cyber-defense.html

References:

[1] "Current Botnet-Techniques and Countermeasures", J. Gassen, E. Gerhards-Padilla, P. Martini. PIK - Praxis der Informationsverarbeitung und Kommunikation. Volume 35, Issue 1, April 2012.

[2] "PDF Scrutinizer: Detecting JavaScript-based Attacks in PDF Documents", F. Schmitt, J. Gassen and E. Gerhards-Padilla. To appear in: Proceedings of the 10th Annual Conference on Privacy, Security and Trust (PST). Paris, France, July 2012.

Please contact:

Jan Gassen, Elmar Gerhards-Padilla
Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Germany
E-mail: jan.gassen@fkie.fraunhofer.de, elmar.gerhards-padilla@fkie.fraunhofer.de

Malware and Botnet Analysis Methodology

by Daniel Plohmann and Elmar Gerhards-Padilla

Malware is responsible for massive economic damage. Being the preferred tool for digital crime, botnets are becoming increasingly sophisticated, using more and more resilient, distributed infrastructures based on peer-to-peer (P2P) protocols. On the other side, current investigation techniques for malware and botnets on a technical level are time-consuming and highly complex. Fraunhofer FKIE is addressing this problem, researching new ways of intelligent process automation and information management for malware analysis in order to minimize the time needed to investigate these threats.

The development and use of malicious software has gained remarkable professionalism in recent years and its impact has drastically expanded. Financially-oriented digital crime operations offer high reward with a low risk of getting caught. Additionally, malware offers huge opportunities in terms of espionage when reviewing the recent incidents of targeted attacks. While attacks are easily carried out through available services and ready-to-go crimeware construction kits, on the defensive side, a thorough analysis of malware can only be conducted by experienced specialists, introducing a strong asymmetry.

The main reason for this is that on the technical side, binary code is the main subject of analysis when in-depth knowledge about malware and botnets is desired. In the analysis process, reverse engineering is the principal technique for granular inspection of functionality and derivation of possible countermeasures. The usual approach is divided into the following stages:

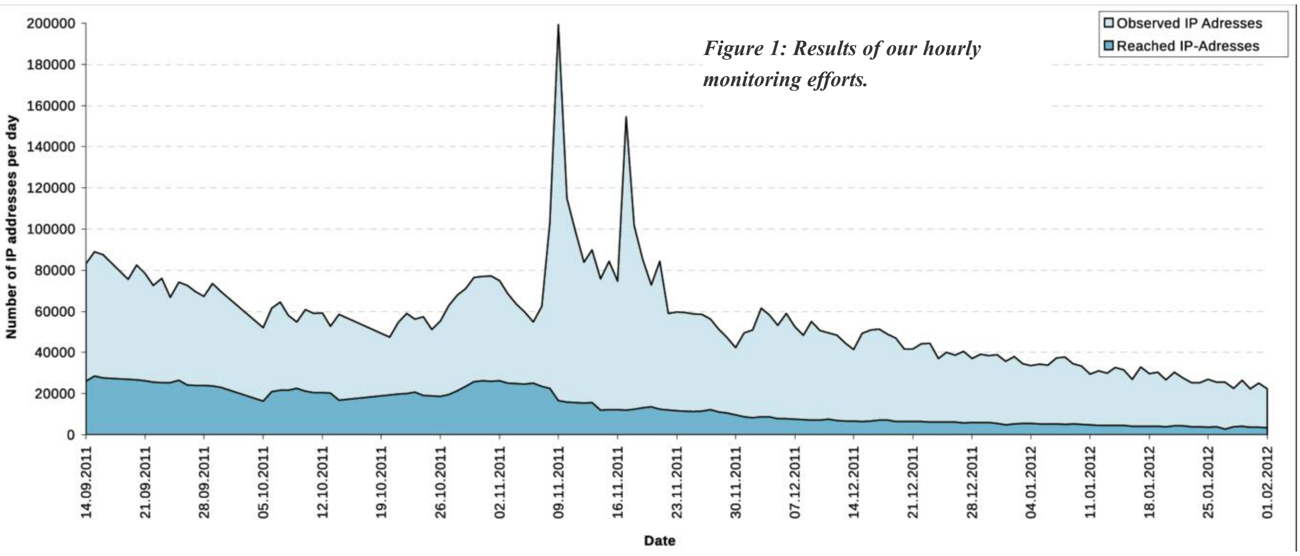


Figure 1: Results of our hourly monitoring efforts.

- A blackbox runtime analysis is performed to gain a first impression of the malware's behaviour.
- Static analysis (no code execution) is applied to the binary to achieve a general structural overview and to identify relevant functional parts.
- Dynamic analysis is performed on these selected parts to extract runtime information that supports static analysis.
- All stages are supported by Open-Source Intelligence (OSINT) methods to link the intermediary technical findings with traces of the malicious operation that are accessible or archived on the Internet.

While the first stage is easily automated through sandboxes, the later stages are mostly performed manually by humans. Reasons for this are protection mechanisms of malware samples, such as anti-reverse engineering techniques and polymorphic representation of semantically-equivalent code. A range of powerful tools cover aspects of the analysis process but they are often neither inter-operable nor suited for immediate collaboration in a team of operators. Furthermore, effective analysis often requires the malware to communicate to its Command & Control (C&C) entity.

Our approach aims at the creation of an analysis environment that bridges the aforementioned gaps. On the one hand we are designing a prototype for an intermediate platform compatible with various tools that allows organization and provision of knowledge extracted from malware samples. On the other hand, we have developed a server component that allows emulation of arbitrary connection endpoints, thus supporting the rapid analysis and adaption of custom communication protocols that is often used by malware.

We have applied parts of the developed process as a test case to investigate the Miner Botnet. This botnet received public attention after starting an extensive Distributed Denial of Service (DDoS) campaign against a range of about 500 German websites, accompanied with attempted extortion for Bitcoins, a virtual currency. The name of the botnet originates from its capability of “Bitcoin Mining”, which is the term used to describe the generation of Bitcoins.

We applied reverse engineering to the set of malicious executables obtained from an initial infection. This allowed us to infer the modular structure of the malware and to extract information about the C&C architecture and derive a specification for the protocol it uses. On these grounds, we created a tracking software for the botnet, iteratively enumerating all active hosts in the P2P layer of C&C architecture.

As an example of findings, Figure 1 shows the results of our hourly monitoring efforts for the period 14 September, 2011 to 1 February, 2012. The figure shows IP addresses that we observed in the circulating address lists and infected hosts we were actually able to reach. The two eye-catching spikes in the observed addresses result from reconfiguration of the botnet by its botmaster, which is not reflected in the actual population. On 1 December, 2011, the last binary update was published to the bots, causing a constantly decreasing trend in the size of the population thereafter. This can be interpreted as the direct natural consequence of the lack of active management of the botnet.

The full results of our investigation of the Miner Botnet were presented in June 2012 at the 4th International Conference on Cyber Conflict (CyCon) in Tallinn, Estonia.

Link:
http://www.fkie.fraunhofer.de/en/research-areas/cyber-defense.html

References:

- D. Plohmann, E. Gerhards-Padilla. “Case Study of the Miner Botnet” in Proc. of the 4th International Conference on Cyber Conflict, Tallinn, Estonia, June 2012.
- D. Plohmann, E. Gerhards-Padilla, F. Leder, "Botnets: Detection, Measurement, Disinfection & Defence", Technical Report published by the European Network and Information Security Agency (ENISA). Editor: Giles Hogben. Heraklion, Greece, March 2011.

Please contact:
Daniel Plohmann,
Elmar Gerhards-Padilla,
Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Germany
E-mail:
daniel.plohmann@fkie.fraunhofer.de,
elmar.gerhards-padilla@fkie.fraunhofer.de

Inference Attacks on Geolocated Data

by Sébastien Gambs

Collecting location data of an individual is one of the greatest offences against privacy. The main objective of this paper is to raise the awareness about the use and collection of such data by illustrating which types of personal information can be inferred from it.

The advent of ubiquitous devices and the growing development of location-based services have lead to the large scale collection of the mobility data of individuals. For instance, the location of an individual can be:

- deduced from the computers IP address,
- collected by applications running on a smartphone providing information tailored to the current location or for collaborative tasks such as traffic monitoring (for example with waze, www.waze.com),
- revealed in form of a geotag, for example added to picture he has taken without him noticing or
- explicitly by checking-in to a geo-social network such as Foursquare (www.foursquare.com).

Among all the Personally Identifiable Information (PII), learning the location of an individual is one of the greatest threats against privacy. In particular, an inference attack, can use mobility data (together with some auxiliary information) to deduce the points of interests characterizing his mobility, to predict his past, current and future locations or even to identify his social network.

A mobility trace of an individual comprises simply a location and a time stamp along with the identifier of the entity behind the trace. From a trail of traces of an individual (ie, a chronological sequence of mobility traces), an inference attack extracts the point of interests (POIs), which correspond to locations frequented by that an individual. A POI could be, for instance, a “home”, a “workplace”, a restaurant visited on a regular basis, a sport centre, a place of worship, the headquarters of a political party or a specialist medical clinic. The semantic of a POI can therefore leak highly personal information about this individual.

The application of a heuristic as simple as finding the location of the last mobility trace before midnight is likely to reveal a person’s “home” and the location whereat the individual spends most of his or her time during the day is likely to be the workplace. In order to identify all of a person’s POIs, a more systematic approach requires the removal of all the traces in which the individual is moving, squashing all the subsequent redundant traces in which the individual stays is immobile for at least 20 minutes into a single point and then running a clustering algorithm on the remaining traces. The output of the clustering algorithm is a set of clusters composed of traces that are close to each others and such that the median of each cluster can be considered as a meaningful POI.

Once the POIs characterizing an individual have been discovered, a mobility model can be constructed. For instance, a mobility Markov chain (MMC), built from the trail of a persons mobility traces, is a probabilistic automaton in which each state corresponds to a POI and an edge indicates a probabilistic transition between two states [1]. The MMC, which represents the mobility behaviour of an individual in a compact and accurate way, can easily be used to predict the next location visited by considering the most probable transition leaving from the current state. Predictions derived from this naïve approach fare 70% to 90% accurate but more sophisticated models can be obtained by remembering the n last visited locations (for *n*=2 or *n*=3) instead of simply the current one or building a MMC for different periods of time (eg, by differentiating between the work days and the week-end or splitting one day into different time slices).

Once a mobility model has been created, it can also be used to perform a de-anonymization attack to identify an individual behind a mobility trace. Suppose, for example, that we have observed Alices movements over a period of time (eg, several days or weeks) during a training phase and that an MMC was derived from her traces. Later, if another geolocated dataset containings mobility traces of Alice is publicly released, the new dataset can be de-anonymized by linking it to the corresponding individuals (Alice) within the training dataset. Simply replacing the names of individuals with pseudonyms before releasing a dataset is rarely sufficient to preserve anonymity because the mobility traces themselves contain information that can be uniquely linked to an individual [2].

Finally, there exists a type of inference attack that partially reconstructs the social graph by assuming that two persons that are regularly in the same neighbourhood at the same time have a high chance of sharing a social link. In the future, it is likely that more and more personal information will be mined from mobility data as the collection of such data increases. For instance, some companies, such as Sense Networks (http://www.sensenetworks.com/macrosense.php), aim to use mobility data to develop detailed profiles of individuals (eg, predicting their income, their social habits or their age) by applying advanced machine learning techniques. The main remaining open question, therefore, is to determine which personal information cannot be inferred from mobility data. In summary, inference attacks highlight the privacy risks raised by the collection of mobility data and show the need to further investigate and design privacy-preserving variants of location-based services providing a fair balance between privacy and the utility of the resulting service [3].

Links:

Creepy:
http://ilektrojohn.github.com/creepy/
MODAP: http://www.modap.org/
Please Rob Me:
http://pleaserobme.com/

References:

[1] S. Gambs, M.-O. Killijian and M. del Prado Cortez, “Show me how you move and I will tell you who you are”, Transactions on Data Privacy 4(2), pp. 103-126, 2011.

[2] P. Golle and K. Partridge, “On the anonymity of home/work location pairs”, Pervasive 2009, pp. 390-397.

[3] M. Damiani, “Third party geolocation services in LBS: privacy requirements and research issues”, Transactions on Data Privacy 4(2), pp. 55-72, 2011.

Please contact:

Sébastien Gambs
Inria – Université de Rennes 1, France
Tel: +33 2 99 84 22 70
E-mail : sebastien.gambs@inria.fr

Secure Enterprise Desktop

by Michael Baentsch, Paolo Scotton and Thomas Gschwind

Using private workstations for business purposes – securely

In recent years, the concept of consumerization has led to consumer IT equipment that is often more powerful and easier to use than workplace computers such as notebooks. This phenomenon, coupled with the rapid generation and renewal cycle of modern ultrabooks and the administrative overhead involved in the purchase of business computers, is fueling the trend of enterprises offering their employees a contribution towards buying a private notebook for use as workplace computer. This is problematic from a security point of view: How can an enterprise ensure that privately owned computers fulfill all security requirements for accessing confidential enterprise data?

Solving the security challenge

The Secure Enterprise Desktop Team (http://www.zurich.ibm.com/secure-ed) of the IBM Research – Zurich Computer Science department believes that it has found an innovative and easy-to-use solution: By using a security token addressing questions of, for example, secure PIN entry, secure back-end connectivity and 2-factor authentication support in an easy-to-use package, namely, the IBM ZTIC (http://www.zurich.ibm.com/ztic) [1], [2] in combination with a secure bootloader and an abstraction layer (hypervisor), a complete separation of private and business-related use of a computer can be implemented: If an employee wants to access data and applications on the enterprise network via a privately-owned computer, he or she must connect the “enterpriseZTIC” to the USB port of the private computer and reboot the latter. The bootloader on the enterpriseZTIC takes control of the computer, establishes a secure connection to the enterprise server, validates the access rights of the employee, and downloads a hypervisor or control oper-

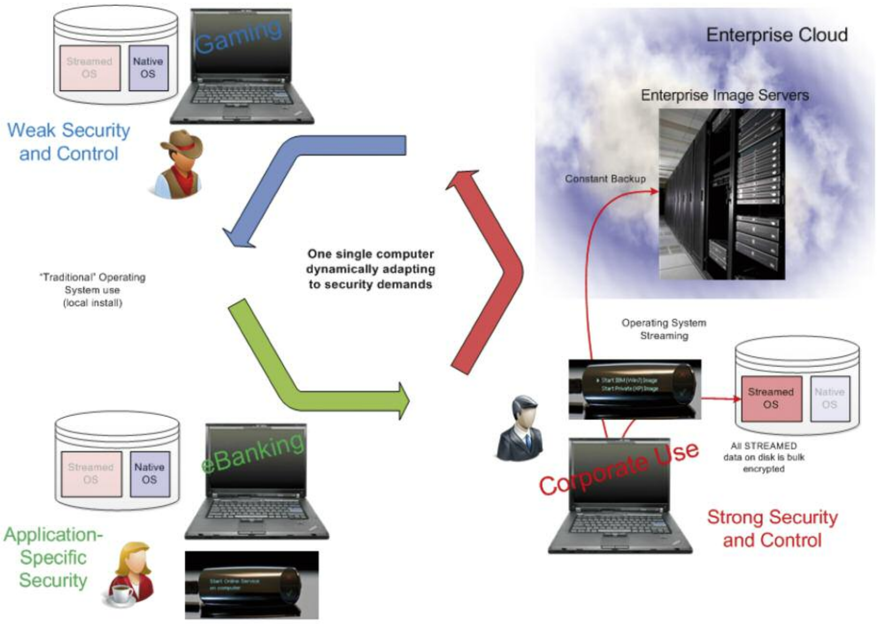

Streamed OS
Native OS
Gaming
Weak Security and Control
"Traditional" Operating System use (local install)
One single computer dynamically adapting to security demands
Streamed OS
Native OS
eBanking
Application-Specific Security
Start Online Service on computer
Enterprise Cloud
Enterprise Image Servers
Constant Backup
Operating System Streaming
Start IBM (Win7) Image
Start Private (XP) Image
Streamed OS
Native OS
All STREAMED data on disk is bulk encrypted
Corporate Use
Strong Security and Control

Figure 1: A single PC with different roles – and security properties – over time.

ating system. As soon as this hypervisor has been installed on the computer, a normal operating system, such as Windows 7® or Linux® can be activated. This normal operating system is provided by the enterprise and contains all security mechanisms necessary for accessing the enterprise network.

The use of the enterpriseZTIC and the secure hypervisor ensures that any malware, for example spy software such as a keyboard logger that may exist in the "private" part of the computer ie the part of the computer that is not "known" to the enterprise, becomes ineffective. This is guaranteed by booting a different operating system that does not utilize the private part of the computer and cannot be modified by any malware present therein. To guarantee fast operation of the entire system, any components of the enterprise operating system that have already been downloaded are encrypted and stored on the hard disk of the computer so that the network does not have to be accessed again to download them. As the secure hypervisor controls all access to the hard disk, the "enterprise operating system" can start up even if some components that are traditionally needed in an operating system have not yet been downloaded. This is achieved by dynamically downloading missing components only when they are requested for the first time.

Advantages and benefits

The system described here offers a number of advantages over a classical solution, such as mere hard-disk encryption to safeguard business secrets: First of all, private use of the computer can readily be permitted as the computer is only used for enterprise purposes in a specially secured software environment. Secondly, thanks to a suitable abstraction layer, the environment of the special "enterprise operating system" becomes very simple, as the abstraction layer will hide any hardware-specific problems by making generic device drivers available. In this way, the tasks normally performed by an IT service centre, such as selection of computers and deployment of customized operating systems, are no longer necessary. Thirdly, the constant synchronization of the hard-disk contents of the computer with the "enterprise operating system" of that particular employee as it is stored on the central server of the enterprise ensures a continual backup of all of the employee's

Figure 2: Overview of the entire system from start up to continual data backup and travel mode.

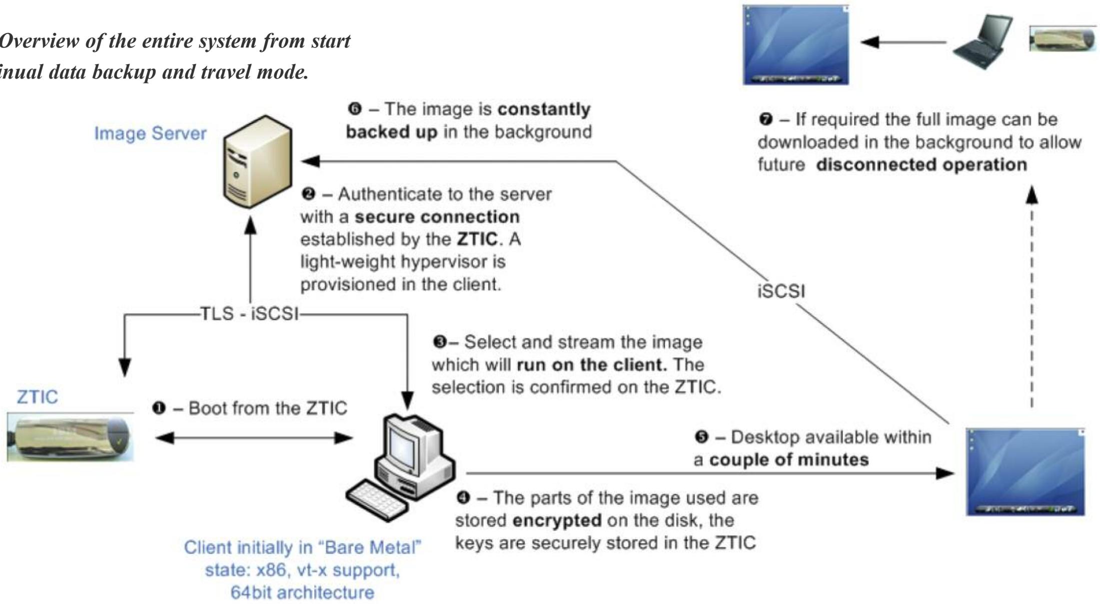


Image Server
❻ – The image is **constantly backed up** in the background
❼ – If required the full image can be downloaded in the background to allow future **disconnected operation**
❷ – Authenticate to the server with a **secure connection** established by the **ZTIC**. A light-weight hypervisor is provisioned in the client.
TLS - iSCSI
iSCSI
❸– Select and stream the image which will **run on the client.** The selection is confirmed on the ZTIC.
ZTIC
❶ – Boot from the ZTIC
❺ – Desktop available within a **couple of minutes**
❹ – The parts of the image used are stored **encrypted** on the disk, the keys are securely stored in the ZTIC
Client initially in "Bare Metal" state: x86, vt-x support, 64bit architecture

data.This also means that the employee can decide to use another computer, start it up with his or her “enterpriseZTIC”, immediately access the enterprise operating system and user data, and thus continue right where he or she left off. In this way, the employee is able to freely switch computers. All data stored on a particular computer are encrypted by the enterpriseZTIC and can only be decrypted and used if that particular enterpriseZTIC is plugged into the USB port of the computer.

A further advantage of this solution is that the images of all enterprise computers are centrally stored and thus can also be centrally managed even if an employee and his or her computer are not physically onsite, ie a corporate campus.

Finally, it is also possible to use this solution in a “travel” mode. To do so, the employee informs the system that he or she wants to start using it without network connection. In this case, all data not yet available locally, ie on the computer, will be downloaded, encrypted and stored. As soon as the system reconnects to the network after offline use, any changes made during offline use will be transmitted back to the server.

Outlook

A special development focus is on solving the challenges involved when computers of one enterprise are used to execute the specifically configured operating system of another enterprise. This situation can arise if third-party employees (external consultants for instance) have to comply with desktop security requirements of the enterprise at which they are temporarily working.

Current state & next steps

The solution described herein is currently deployed as internal pilot, and the development team is actively searching suitable first users outside of IBM.

Links:
http://www.zurich.ibm.com/secure-ed
http://www.zurich.ibm.com/ztic

References:

[1] T. Weigold, et.al.: The Zurich Trusted Information Channel – An Efficient Defence against Man-in-the-Middle and Malicious Software Attacks; In P. Lipp, A.-R. Sadeghi, and K.-M. Koch (Eds.): TRUST 2008, LNCS 4968, pp. 75–91, 2008. © Springer-Verlag Berlin Heidelberg 2008.

[2] M. Baentsch, et.al.: A Banking Server's Display on your Key Chain; ERCIM News 73, 2008; http://ercim-news.ercim.eu/content/view/345/543/

Please contact:

Michael Baentsch
IBM Research Zurich
Tel: +41 44 724 8620
E-mail: secure-ed@zurich.ibm.com

Advances in Hash Function Cryptanalysis

by Marc Stevens

When significant weaknesses are found in cryptographic primitives on which the everyday security of the Internet relies, it is important that they are replaced by more secure alternatives, even if the weaknesses are only theoretical. This is clearly emphasized by our construction of a (purposely crippled) rogue Certification Authority (CA) in 2009 that in principle enabled the impersonation of all secure websites. This was possible due to the continued use of the insecure cryptographic hash function MD5 by a leading commercial CA. The hash function SHA-1, the successor to MD5 as the de facto hash function standard, has been theoretically broken since 2005. The Cryptology group at CWI has recently made a significant step towards a practical attack on SHA-1 that has long been anticipated, as well as efficient counter-measures against these cryptographic attacks.

Cryptographic hash functions, such as MD5, SHA-1 and SHA-2-256, are among the most important cryptographic primitives. A hash function is an algorithm that computes a hash value of a fixed number of bits (say 256 bits) for a message of arbitrary bit-length. A main application of hash functions is in digital signatures. In order for digital signatures to be secure, a hash function must satisfy the collision resistance property: it must be hard to find collisions, i.e., two different messages that map to the same hash value.

In 2004, collisions were found for MD5 by Wang et al. and despite practical limitations, MD5 was found to be insecure for continued use in applications. We have introduced the chosen-prefix collision attack in 2006 that removes limitations of the identical-prefix collision attack and thereby results in significantly more potential for realistic threats to the security of digital signatures. In particular, due to the slow response of the industry in removing MD5, we were able to construct a rogue Certification Authority (CA) in 2009. The certificate of our rogue CA was signed by an unsuspecting commercial CA. We used an improved version of the MD5 chosen-prefix collision attack to do this, thereby effectively breaking the security of secure websites (https://) [1].

A similar situation exists for SHA-1 now as for MD5 in 2005. It has been theoretically broken since 2005 due to a collision attack presented by Wang et al. with a complexity of 2^{69} SHA-1 computations. Since then there have been several claims of improved attacks with complexities as low as 2^{52} SHA-1 computations, however these were either not substantiated to date, withdrawn, or found to be too optimistic. Unfortunately, this means that the first attack essentially remains the state of the art in the literature. Lack of recent improvements suggests a barrier has been reached.

Recently we have introduced a new exact cryptanalysis of SHA-1 that, as prior methods, is based on the approach of local collisions from which an appropriate system of equations is obtained,

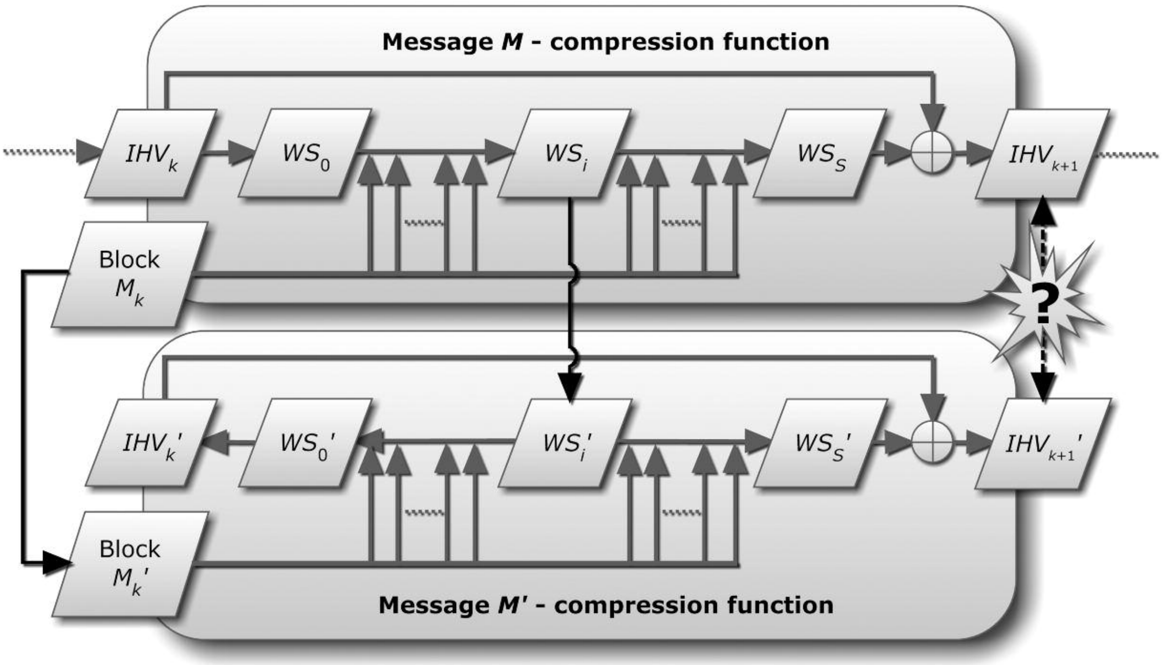


Figure 1: Detection of whether a message has been constructed using a collision attack on the cryptographic hash functions MD5 and/or SHA-1. This is done by partially reconstructing the hash computation of the (unknown) colliding sibling message (bottom half) and looking for the tell-tale condition of a collision (the comparison on the right).

which is then subsequently used in the search for an actual collision [2]. The novelty of our new methods is two-fold. First, in theory there are many eligible appropriate systems one may arrive at in this approach. Our analysis of such systems is exact and does not use heuristics compared to prior methods, in particular with respect to the dependence of local collisions and determining the complexity. Secondly, we show, for the first time, how to efficiently pick the system of equations that leads to the lowest complexity among all considered systems for a particular combination of local collisions. This is achieved by analyzing partitions of the set of all possible so called exact differential paths and using inherent redundancies to be able to compute the exact total probability of each partition. These probabilities are used to derive the optimal system of equations.

With our new identical-prefix collision attack based on our new cryptanalytic methods, we have shown how to substantially reduce the complexity of finding collisions for SHA-1 to about 2^{61} SHA-1 computations. Though this is still just out of reach, this is a preliminary attack based on our new methods and the attack implementation can be further improved. The implementation of our SHA-1 attack is the first to be publicly available, thus its correctness and complexity can be publicly verified and also allows further understanding and improvements [3].

As MD5 and SHA-1 have significant (theoretical) weaknesses, they evidently should be withdrawn from applications. However, practice shows that the industry responds slowly in replacing them with secure hash function standards. To alleviate possible damage by collision attacks, we have introduced a technique that efficiently detects both identical-prefix and chosen-prefix collision attacks against both MD5 and SHA-1 given only one of the two documents in a collision. Such an indication can be used to abort further processing or communications, before sensitive information can be accessed or transmitted.

The future de facto hash function standard SHA-3 is currently being selected in an international competition by the National Institute of Standards and Technology (NIST) in the U.S. Nevertheless, due to the continued usage of SHA-1 in the foreseeable future, more research is needed on the real-world security of SHA-1 and on whether our ideas can be extended to other important hash function standards such as the future SHA-3.

Links:
http://www.cwi.nl/crypto
http://marc-stevens.nl/research

References:

[1] M. Stevens, A. Sotirov, J. Appelbaum, A. Lenstra, D. Molnar, D. A. Osvik, B. de Weger, "Short Chosen-Prefix Collisions for MD5 and the Creation of a Rogue CA Certificate", CRYPTO 2009, Lecture Notes in Computer Science, vol. 5677, Springer, 2009, pp. 55-69.
http://marc-stevens.nl/research/papers/CR09-SSALMOdW.pdf

[2] M. Stevens, "Attacks on Hash Functions and Applications", PhD thesis, Leiden University, June 19, 2012.
http://www.cwi.nl/system/files/PhD-Thesis-Marc-Stevens-Attacks-on-Hash-Functions-and-Applications.pdf

[3] Project HashClash, Marc Stevens, http://code.google.com/p/hashclash.

Please contact:

Marc Stevens, Cryptology Group, CWI, The Netherlands
Tel: +31 20 592 4195
E-mail: Marc.Stevens@cwi.nl

Social Snapshot Framework: Crime Investigation on Online Social Networks

by Markus Huber

Recently, academia and law enforcement alike have shown a strong demand for data that is collected from online social networks. We present a novel method for harvesting such data from social networking websites. Our approach uses a hybrid system based on a custom add-on for social networks in combination with a web crawling component.

Over recent years, online social networks (OSNs) have become the largest and fastest growing websites on the Internet. OSNs, such as Facebook or LinkedIn, contain sensitive and personal data of hundreds of millions of people, and are integrated into millions of other websites. Online social networks continue to replace traditional means of digital storage, sharing, and communication. Collecting this type of data is thus an important problem in the area of digital forensics. While traditional digital forensics is based on the analysis of file systems, captured network traffic or log files, new approaches for extracting data from social networks or cloud services are needed. Despite the growing importance of data from OSNs for research, current state of the art methods for data extraction seem to be mainly based on custom web crawlers.

Our approach is based on a hybrid system that uses an automated web browser in combination with an OSN third-party application. Our system can be used efficiently to gather ``social snapshots", datasets which include user data and related information from the social network. The datasets that our tool collects contain profile information (user data, private messages, photos, etc.) and associated meta-data (internal timestamps and unique identifiers). We implemented a prototype for Facebook and evaluated our system on a number of volunteers.

Figure 1 shows the core applications of our social snapshot framework. (1) The social snapshot client is initialized by providing the target user's credentials or cookie. Our tool then starts the automated browser with the given authentication mechanism. (2) The automated browser adds our social snapshot application to the target user's profile and sends the shared API secret to our application server. (3) The social snapshot application responds with the target's contact list. (4) The automated web browser requests specific web pages of the user's profile and her contact list. (5) The received crawler data is parsed and stored. (6) While the automated browser requests specific web pages our social snapshot application gathers personal information via the OSN API. (7) Finally the social data collected via the third-party application are stored on the social snapshot application server.

In order to get access to the complete content of a target’s social network account, social snapshots depend on gathering the initial authentication token. Below we outline three digital forensic scenarios, representative of real-world use cases, which illustrate the initial gathering of the authentication token.

Consent: This naive approach requires consent from the person whose social networking profiles are analysed. A person would provide the forensic investigator temporary access to her social networking account in order to create a snapshot. This would also be the preferred method for academic studies to conduct this research ethically. and to comply with data privacy laws.

Hijack social networking sessions: Our social snapshot application provides a module to hijack established social networking sessions. An investigator would monitor the target’s network con-

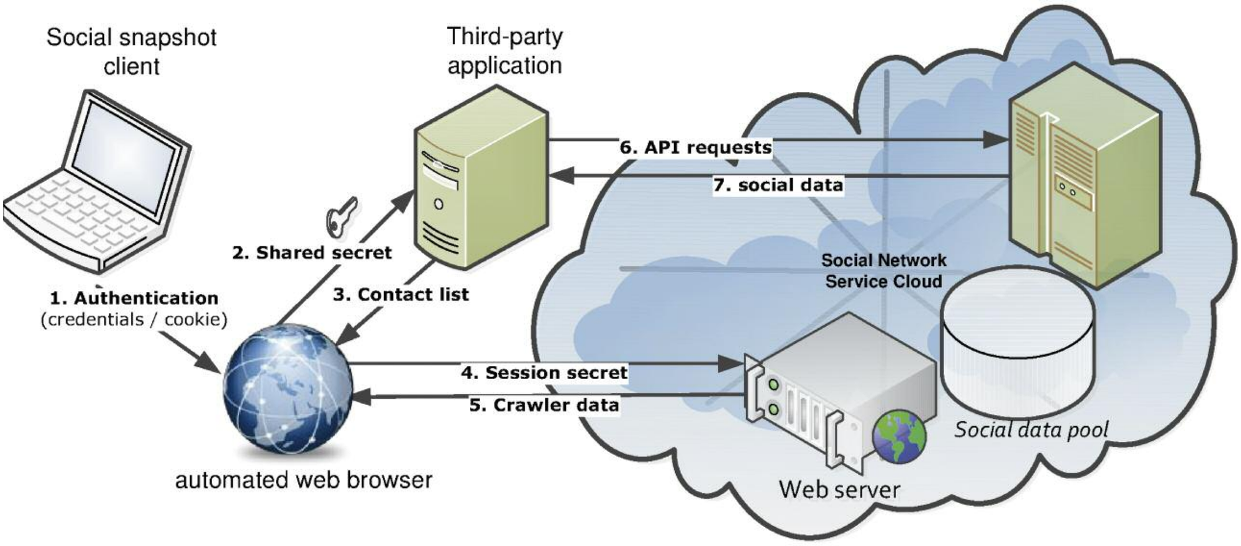


Figure 1: Collection of digital evidence through our social snapshot framework

nection for valid authentication tokens, for example unencrypted WiFi connections or LANs. Once the hijack module finds a valid authentication token, the social snapshot application spawns a separate session to snapshot the target user's account.

Extraction from forensic image: Finally, physical access to the target's personal computer could be used to extract valid authentication cookies from webbrowsers. Stored authentication cookies can be automatically found searching a gathered hard drive image or live analysis techniques.

Social snapshots explore novel techniques for automated collection of digital evidence from social networking services. Compared with state-of-the-art web crawling techniques our approach significantly reduces network traffic, is easier to maintain, and has access to additional and hidden information. Extensive evaluation of our techniques has shown that they are practical and effective in collecting the complete information of a given social networking account reasonably fast and without detection from social networking providers. We believe that our techniques can be used in cases where no legal cooperation with social networking providers exists. In order to provide a digital evidence collection tool for modern forensic investigations of social networking activities, we release our core social snapshot framework as open source software. We will continue to extend the analysis capabilities of our forensic software and cooperate with partners on the evaluation of real-world cases.

The research was funded by COMET K1, FFG - Austrian Research Promotion Agency, by the Austrian Research Promotion Agency under grants: 820854, 824709, and 825747. Further details of this work can be found in our previously published paper [1].

Links:
http://www.sba-research.org/socialsnapshots
https://github.com/markushuber/social-snapshot-tool

References:

[1] Markus Huber, Martin Mulazzani, Manuel Leithner, Sebastian Schrittwieser, Gilbert Wondracek, and Edgar Weippl. 2011. Social snapshots: digital forensics for online social networks. In Proceedings of the 27th Annual Computer Security Applications Conference (ACSAC '11). ACM, New York, NY, USA, 113-122. http://www.sba-research.org/wp-content/uploads/publications/social_snapshots_preprint.pdf

Please contact:

Markus Huber
SBA Research (AARIT), Austria
E-mail: mhuber@sba-research.org

TorScan: Deanonymizing Connections Using Topology Leaks

by Alex Biryukov, Ivan Pustogarov and Ralf-Philipp Weinmann

Tor is one of the most widely used tools for providing anonymity on the Internet. We have devised novel attacks against the Tor network that can compromise the anonymity of users accessing services that exhibit frequent and predictable communication patterns and users establishing long-lived connections.

Anonymity on the Internet is a double-edged sword. On the one hand, anonymity allows people to express their thoughts and ideas without fear of repression; on the other hand it can be used to commit crimes in the digital domain with impunity. The Tor network is one of the most popular and widely used tools for enabling anonymous Internet communications. By routing connections through a variable chain of three "relays", volunteer-operated Tor servers, the origin of the user establishing the "circuit" is cloaked.

As a consequence, services that the user connects to do not see the real IP address of the user but rather the IP address of the last computer in the chain of relays. To make this mechanism more secure, the series of servers used for new connections is changed every ten minutes and connections between Tor relays multiplex sessions of multiple users. Moreover, the first relay in the chain, the so-called "guard", is picked from a small set of relays – usually consisting of only three elements – that is randomly chosen by the user's Tor client on its first start. The concept of guard nodes was introduced to give users a chance to avoid falling prey to attackers controlling a fraction of the Tor network. Guard relays remain in this set for approximately one month.

We have found techniques to scan the connectivity of Tor relays. These can provide us with a topological map of the Tor network (see Figure 2 for an example).

With a speed of approximately 20 seconds per scanned router or under three minutes for scanning the whole network in a parallelized manner we are also able to observe dynamics in the topology.

Based on these data we have devised attacks that compromise the anonymity of users exhibiting one of the following two communication patterns over the Tor network:

- *Long-lived connections:* Large downloads, a number of instant messaging protocols and BitTorrent over Tor will create circuits that are kept alive for many hours. Looking at differences in the topology to identify connections between relays that remain stable enables us to trace the communication back to the guard node.
- *Frequent, identifiable reconnections:* Some Internet services, for example webmail services such as GMail, will

frequently re-establish connections to a server. In case these connections can be identified to belong to the same user, an attacker that can observe a good chunk of the exit traffic of the Tor network is also able to trace the user back to a set of guard nodes. In particular, we will be able to identify guard nodes with low bandwidth.

As the Tor network steadily grows in size, we have created a model to estimate the effectiveness of our attacks for arbitrary parameters. The main parameter affecting the connectivity of a relay is the bandwidth it provides to the Tor network. The probability of any given relay being chosen for a route mainly hinges on the bandwidth contributed. Even though there are other parameters influencing this probability, it only takes a very simple model to predict the observed connectivity well.

Experimental verification of our attacks against connections made by ourselves over the real Tor network show that they work in practice and our model predicts their effectiveness correctly.

All prior research on Tor assumed the topology of the Tor network to be opaque. Since our attack methodology is orthogonal to other attacks described in the literature, inference of the Tor network topology can also be used to enhance these attacks.

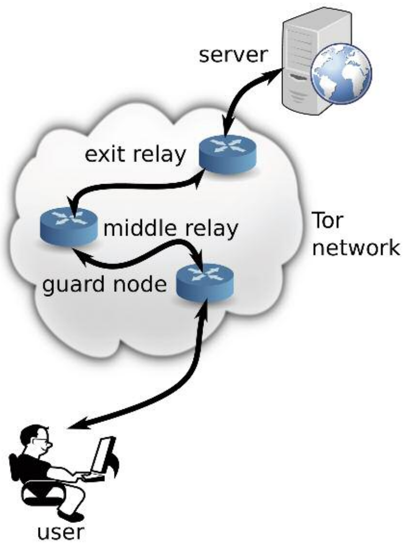


Figure 1: User contacts server through the Tor network.

The work described in this article has been accepted to the European Symposium on Research in Computer Security (ESORICS 2012) and the Tor Project has been made aware of our findings. In the paper [2] we also describe countermeasures that can be used to mitigate our attacks.

Links:
Tor Project: https://www.torproject.org
CryptoLUX group:
https://www.cryptolux.org
Interdisciplinary Centre for Security, Reliability and Trust:
http://www.uni.lu/snt

References:

[1] R. Dingledine, N. Mathewson, P. Syverson: Tor: The Second-Generation Onion Router in Proceedings of the 13th USENIX Security Symposium, August 2004, p. 303-320

[2] A. Biryukov, I. Pustogarov, R.-P. Weinmann: TorScan: Tracing Long-lived Connections and Differential Scanning Attacks, Proceedings of ESORICS 2012, to appear

Please contact:

Ivan Pustogarov
University of Luxembourg,
Luxembourg
E-mail: ivan.pustogarov@uni.lu

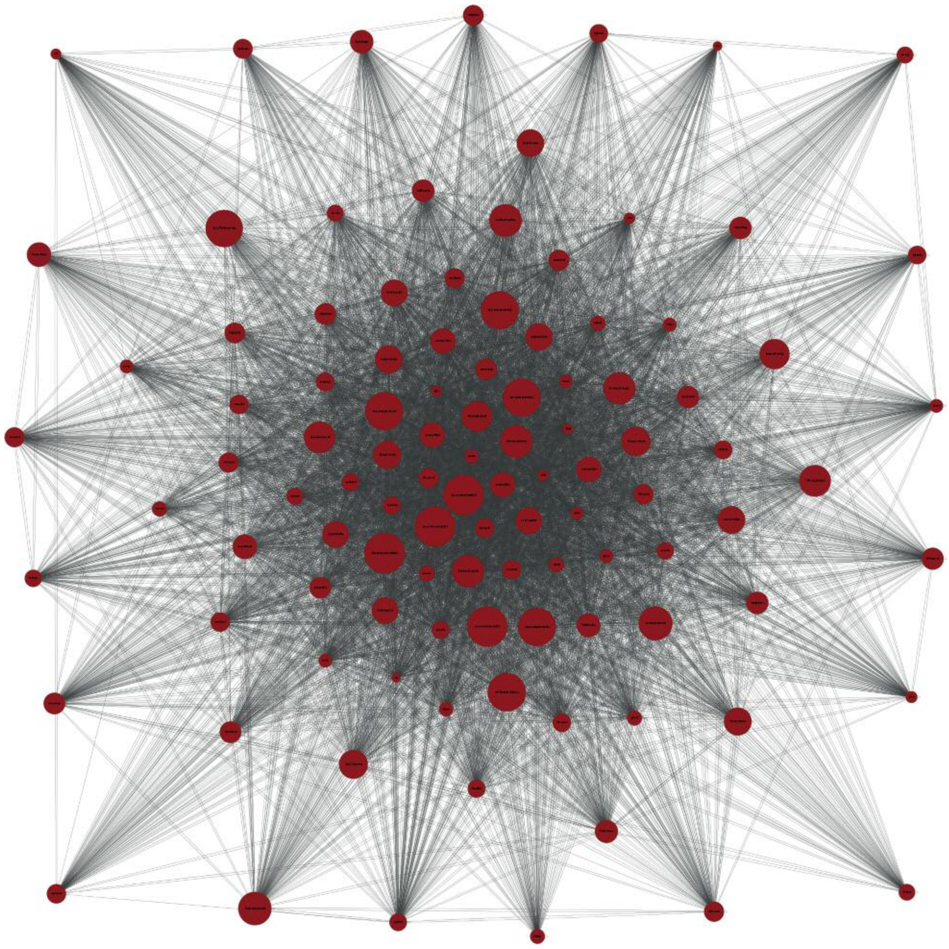


Figure 2: Connectivity of the 105 fastest Tor relays on March 14th, 2012, 13:59 GMT.

Visualization for Monitoring Network Security Events

by Christopher Humphries, Nicolas Prigent and Christophe Bidan

As networks increase in size and complexity, IT security officers are being overwhelmed by large volumes of data (alerts from IDSes, logs from various machines and services, etc). These data are often very heterogeneous and multidimensional. It is, of course, impossible to handle these data manually and even automated analysis tools are often inadequate owing to the scale of the data. This situation has reached a point at which most available data is never used. Visualization provides new hope in this context.

Visualization provides a means for representing and understanding large amounts of data in a synthetic and often aesthetic way, allowing operators to handle large amounts of data more easily. In the context of network security event monitoring, visualization provides benefits for the following activities:

- Realtime supervision, in which visualization tools emphasize and prioritize malicious events that have been detected. They also provide "situational awareness", ie general information on what is happening in the system. In this case, visualization rarely provides detail but instead offers a global view of the system.
- Forensic analysis, in which visualization tools allow operators to mine various data sources for details about what happened in the system. Visual data mining tools are often used on a subset of data, this subset having been selected based on the detection of malicious and/or abnormal activities.
- Fast reaction, in which visualization tools help the operator to quickly react when faced with security events, eg by re-configuring a firewall or automatically deploying a patch on vulnerable systems. Visualization for fast reaction often provides supplementary information about the state of the network and security policies.
- Communication, in which visualization tools help operators explain situations with more clarity. The generated representations can be for internal purposes (enhanced ticketing systems for instance) or for external purpose, ie to improve communications about a situation a posteriori.

Many visualization tools can be used for network security event monitoring. Some are very generic and offer representations similar to those made available in traditional spreadsheet software (bar charts, pie charts, radar graphs, tree maps, etc). Others are very specific, taking specific data formats into account to offer more sophisticated representations, designed with very specific objectives in mind. For example, The Network Visualizer (TNV) [1] uses packet capture (pcap) files to represent network communication flows.

Consequently, operators in charge of monitoring the networks or the analysts performing forensic tasks are bound either to use very specific tools and hope that they will fit their every need, or to use very generic tools and sacrifice precision and context. To perform well, the latter solution requires that the user be skilled not only in network security but also in data analysis and visual information design. In fact, finding the right visualization for a given set of data requires choosing the best visual representation for the data context and current objectives. Although experts have at least partially addressed this issue [2], choosing the correct representation still requires a lot of time and experience for security specialists, especially when a background in visualization or statistics is lacking.

The security visualization workgroup in the CIDre team is currently working on a system that assists security specialists in handling and exploring security-related data. Our goal is to allow users to monitor and explore their data in a way that is as user-friendly as possible. Ideally, the user does not need to be a specialist in design or visualization: One of our long term objectives is to enable the system to automatically generate the adequate representations according to the current data and visual contexts, guided by the goals and intentions of the user. He or she is only aware of the data sources that are available and interacts with them in a natural way. Therefore, details about the requests that are made to the datasets have to be hidden. To this end, we are working on the automated translations of goals and intents from the user into traditional database requests, though not necessarily SQL. Our system also automatically selects the representations that are best suited to the user's objectives.

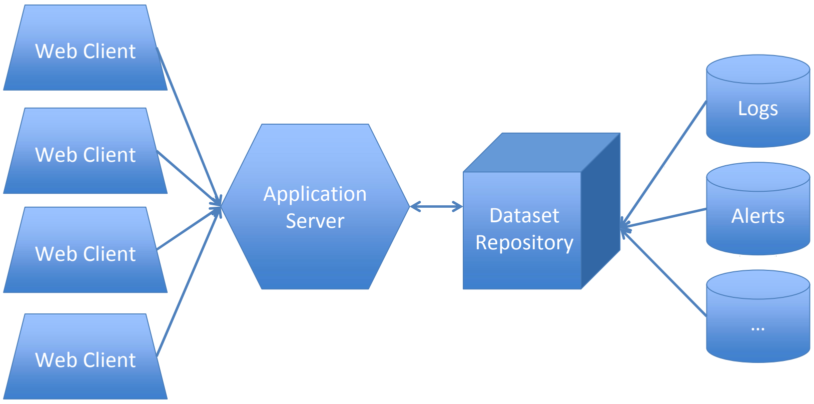


Figure 1: Architecture of the visualization system

Finally, since forensic science is by its very nature an interactive process, we are working on the interactions between user and data through representation dynamics.

While efficient visualization is often a user-centric problem, building a responsive and performant system is essential when dealing with large volumes of security-related data. To take this aspect into account, we are currently working on the following web-oriented architecture (see Figure 1): The dataset repository collects, stores, indexes and serves the various required datasets (snort alerts, log files, etc.). An application server accesses this dataset repository and serves the client web application and assets to multiple endpoint clients. It acts as a data server proxy and in so doing provides midpoint data caching and more powerful mathematical and statistical operations for the client. The web application itself has network components for data requests and persistent connections, and has similar caching and data processing abilities. The final layer, after the main application logic, holds our visualization components for representing and interacting with data.

By proposing new ways to interact with security-related data as well as an efficient architecture to do so, we hope to provide a more efficient, portable and flexible option for visualizing, exploring and monitoring network security events.

Link:
[1] http://tnv.sourceforge.net/

Reference:

[2] Leland Wilkinson and Graham Wills, The Grammar Of Graphics, 2nd edition, Springer, 2005

Please contact:

Nicolas Prigent
SUPELEC, France
E-mail: nicolas.prigent@supelec.fr

Challenges of Cloud Forensics: A Survey of the Missing Capabilities

by Rafael Accorsi and Keyun Ruan

‘Cloud computing environments have become a new battlefield for cyber crime’ [3], posing an unprecedented risk to individuals’ privacy. A survey with 257 respondents on Cloud forensic capabilities and perceived challenges shows the state of Cloud forensics.

Cloud computing is radically changing the way information technology services are created, delivered, accessed and managed, as well as the corresponding business models. Keeping pace with the continuous adoption of Cloud computing, a rapidly increasing amount of data is being transferred to and processed by clouds. Trusting massive quantities of data processing to third parties raises security and privacy concerns, which – despite existing mechanisms – turn out to be the main inhibitor for Cloud adoption.

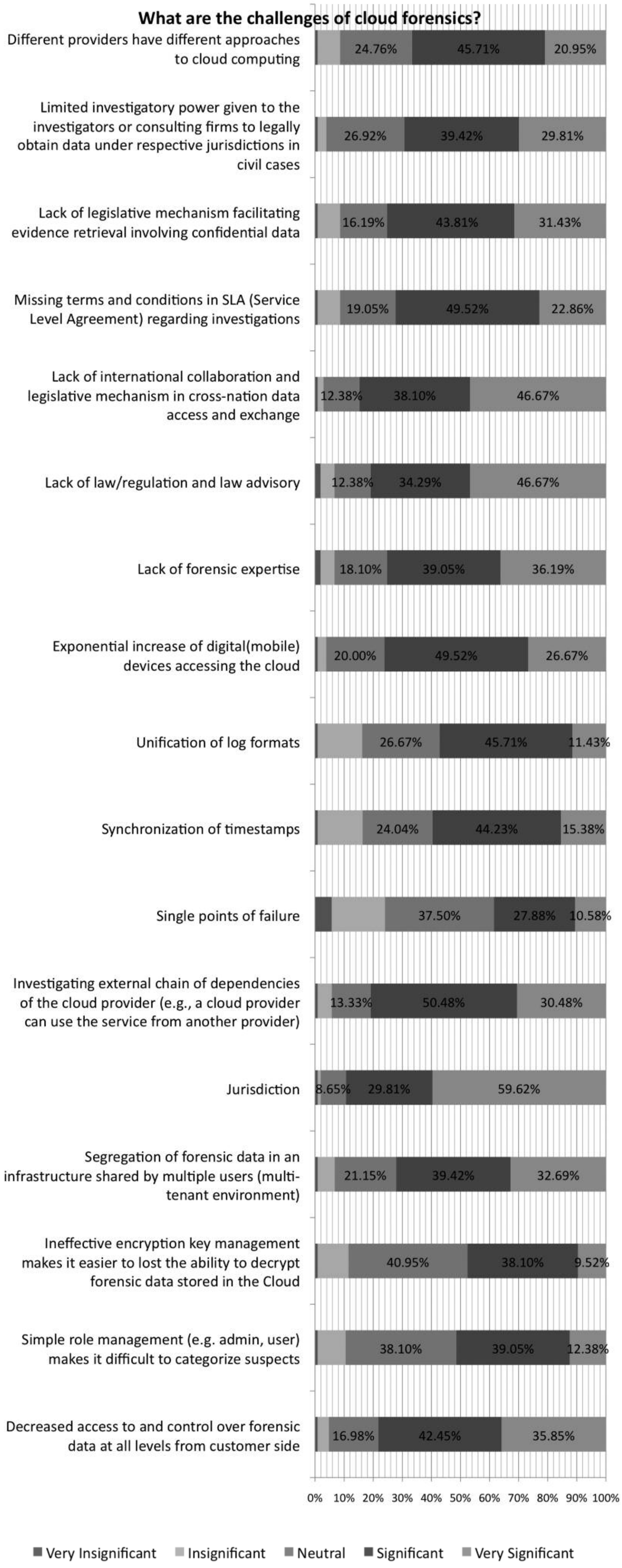
If available, forensic techniques could be used to provide accountability and hence re-establish trust. A representative survey conducted by researchers [3] captured the current state of “Cloud forensics”, ie the application of digital forensics in Cloud computing environments. The survey is a primer in this research field. It received 257 responses worldwide, from academics, practitioners, service providers and industry. The top challenges for Cloud forensics are depicted in Figure 1. The complete survey is under review at Elsevier’s Digital Investigation; a preliminary status can be found in [2].

The background part of the survey regards the definitions of and expectations of Clouds. The majority of the respondents (62%) agree that the Cloud is not a new technology but a novel means of delivering computing resources. This does not imply that all the technical means needed to provide reliable Cloud is at hand: in fact 70% of the respondents disagree with this. Over 49% of participants believe that Cloud computing reduces cost and compromises security. Some argue that the rapid growth of Cloud computing is driven by cost reduction with known risk and sacrifice of security. This popular argument among early-adopters is only partly supported by the respondents of this question, but backed in other surveys carried out in the US and Europe, where up to 70% of the participants are concerned with security.

The survey highlights the importance of forensic techniques tailored to the Cloud: 81% of respondents agree that forensics is an indispensible component of Cloud security, while 76% claim that this area needs more funding and investment than it currently receives. Interestingly, 71% of respondents believe that the general lack of awareness of Cloud security will endure until a major critical incident happens. This could explain the aforementioned fact that security concerns are not levelled as “critical” by respondents. The results of this question show that the respondents have reached consensus on the significance of Cloud forensics. However, leading organizations driving Cloud security standards still largely neglect the importance of integrating forensic capabilities into Cloud security in their most recent releases.

Up to 41% of participants believe that Cloud forensics is a “brand new area” and only 25% agree that Cloud forensics “is” classical computer forensics”. Up to 46% of the participants believe that Cloud computing makes forensics harder. The reasons are manifold, including reduced access the physical infrastructure and storage (26%), lack of standard interfaces (15%) and evidence segregation (12%). Among the 37% believing that Cloud makes forensics easier, 17% state that evidence is harder to destroy as a result of mirroring, 15% consider that forensic functionalities can be integrated into Clouds and 10% believe that there will be less work for the investigator/law enforcement side because the data are centralized on the provider.

Regarding the focus of Cloud forensics, 56% of respondents regard it as an inter-



disciplinary effort, in contrast to 44% believing it to be purely technical. 80% of respondents agree that there is a "technical" as well as "legal" dimension for Cloud forensics; 14% of respondents clicked "other" dimensions. "Political" and "personal" dimensions are mentioned in the comments. There was a consensus among respondents that Cloud forensics comprises three major dimensions: technical, organizational and legal.

As for forensics usage, 80% of respondents agree that Cloud forensics can be used for "investigations on digital crimes, civil cases, policy violations, etc.", 51% agree that it can be used for "regulatory compliance", 46% agree it can be used for "data and system recovery", 40% agree that it can be used for "due diligence" and 34% agree that it can be used for "log monitoring". Since Cloud forensics is an application of digital forensics in Cloud computing, its usage should be similar to the usage of digital forensics in general. When applied in Cloud computing environments, the split of control among Cloud actors has made forensics a shared responsibility which adds to the organizational complexity of Cloud forensics.

Taking stock, the survey results show that Cloud adoption does pose significant novel challenges to digital investigation, rather than scaling up existing problems. In fact, there is to-date a lack of mechanisms to address forensic investigations in the Cloud [1,2], as well as solid jurisdiction on handling Cloud-related cases [2]. This gives incentives to defectors exercising cybercrime. Researchers strongly believe that it is a critical timing for standard acceleration and enhancing forensic capabilities while the technology matures.

Links:
BPSec Group: http://bpsec.telematik.uni-freiburg.de/
Cloud Forensics Research:
http://www.cloudforensicsresearch.org/

References:

[1] Accorsi, R.; Wonnemann, C. (2011): Forensic leak detection for business process models. Advances in Digital Forensics VII, pages 101-113. Springer.
[2] Ruan, K.; Carthy, J.; Kechadi, T.; Crosbie, M. (2011): Cloud forensics. Advances in Digital Forensics VII, pages 35-46. Springer.
[3] Ruan, K.; Baggili, I.; Carthy, J.; Kechadi, T. (2011): Survey on cloud forensics and critical criteria for cloud forensic capability: A preliminary analysis. ADFSL Conference on Digital Forensics, Security and Law.

Please contact:

Rafael Accorsi
University of Freiburg, Germany
E-mail: rafael.accorsi@iig.uni-freiburg.de

Keyun Ruan
Centre for Cybersecurity and Cybercrime Investigation
University College Dublin, Ireland
E-mail: keyun.ruan@ucd.ie

Domain-Specific Languages for Better Forensic Software

by Jeroen van den Bos and Tijs van der Storm

Recovering evidence of criminal activities from digital devices is often difficult, time-consuming and prone to errors. The Software Analysis and Transformation group at CWI designed Derric, a domain-specific language (DSL) that enables the efficient construction of scalable digital forensics tools.

An important part of digital forensics is recovering evidence from digital devices. This typically includes recovery of images, text documents and email messages relevant to a forensic investigation. Currently, such investigations crucially depend on custom-made software, which often has to be modified on a case-by-case basis. Additionally, it needs to scale to deal with datasets in the terabyte range. CWI applies state-of-the-art language engineering tools and techniques to make the construction and maintenance of such software less error-prone and time consuming.

An application area for digital forensics software is “file carving”, the process of recovering files from a digital device without the use of file system metadata. File carving is used, for instance, to recover child pornography images, even though the suspect may have tried to delete them. Moreover, because of fragmentation, a file may be distributed over a device in multiple fragments. File carvers then match sequences of bytes to be of a certain file format and attempt to reconstruct the original file.

File formats, such as JPEG (images), ZIP (archives) and DOC (documents), play a crucial role in file carving. They define the structure necessary to determine if a raw file fragment might be part of a complete file of a certain type. File formats exist in many versions and vendor-specific variants. In the current state of practice, file format knowledge is often intertwined with complex, highly optimized, file carving algorithms for reassembling parts of fragmented files. This lack of “separation of concerns” makes changing forensic software error-prone and time consuming.

Derric: a DSL for file formats

Derric,is a domain-specific language (DSL) designed by CWI that can be used to describe file formats. Such descriptions are then input to a code generator that generates high-performance file carvers. This way knowledge about file formats is isolated from the algorithmic file carving code. Forensic investigators can focus on maintaining and evolving file format descriptions, whereas software engineers can focus on optimization of the runtime system.

A Derric description consists of three parts: a configuration header, the sequence section, and a list of structure definitions. The configuration header declares file type metadata, such as endianness, signedness and string encodings. The sequence section then describes the high-level structure of a file using a regular expression. Finally, the tokens used in the regular expression are defined in the structure section. Each structure is identified by a name and contains one or more fields. The contents and length of a field may be arbitrarily constrained in order to guide the matching process. In our experience, Derric is expressive enough to describe a wide range of file formats.

We have evaluated Derric by comparing generated file carvers to existing file carvers that are used in forensic practice [1]. Our results show that the Derric-based file carvers perform as well as the best file carvers out there, and sometimes even better. Derric is implemented in Rascal, a metaprogramming language and its implementation is very small: around 2000 lines of Rascal, and a runtime library of 4200 lines of Java Code. As a result, the overhead of maintaining the DSL implementation is acceptable.

An additional advantage of declaratively describing file formats using Derric is that the descriptions can be transformed before passing them to the code generator. Source-to-source transformation can be applied to configure the trade-off between runtime performance and accuracy. We have implemented three such transformations for successively obtaining carvers that are more efficient. In certain forensics cases, it may be more effective to compromise on accuracy in order to obtain results more quickly. Since the transformations are fully automated this trade-off can be made without having to change any code. We have evaluated the effect of the transformations on a 1TB test image. Our results show that performance gains up to a factor of three can be achieved, at the expense of up to 8% in precision and 5% in recall.

Conclusion

Digital forensics, now more than ever, is crucially dependent on software. DSLs can help untangle the concerns that are at play in the domain of forensics. Derric is an important step in this direction: by separating file format descriptions from how they are used in implementation, forensic tools become easier to modify. Moreover, model transformation provides opportunities for configuring trade-offs that would otherwise be cast in stone.

Links:

Derric: http://www.derric-lang.org/
Rascal: http://www.rascal-mpl.org

References:

[1] J. van den Bos, T. van der Storm, “Bringing Domain-Specific Languages to Digital Forensics”, in: Proc. of the 33rd International Conference on Software Engineering (ICSE'11), Software Engineering in Practice, ACM, 2011

[2] J. van den Bos and T. van der Storm, “Domain-Specific Optimization in Digital Forensics”, in: Proc. of the 5th International Conference on Model Transformation (ICMT'12), 2012

Please contact:

Jeroen van den Bos, Tijs van der Storm
CWI, The Netherlands
E-mail: J.van.den.Bos@cwi.nl,
T.van.der.Storm@cwi.nl

Legal Issues Associated with Data Management in European Clouds

by Attila Kertesz and Szilvia Varadi

Cloud Computing offers flexible resource provision for businesses, enabling them to respond effectively to new demands from customers. This new technology moves local data management to a third-party provided service, a phenomenon that raises legal issues such as data protection and privacy. We have evaluated Cloud use cases against the applicable law set out by the Data Protection Directive of the EU to pinpoint where legal problems may arise.

Cloud Computing offers on-demand access to infrastructure resources operated from a remote source. Recently, this form of service provision has become hugely popular, with many businesses migrating their IT applications and data to the Cloud to take advantage of the flexible resource provision that can benefit businesses by responding quickly to new demands from customers. Cloud Computing also moves functions and responsibilities away from local ownership and data management to a third-party provided service, and brings with it a set of associated legal issues such as data protection and privacy, and the need to comply with certain regulations. Owing to the pace of technical and economic progress in this field it is important to determine the compliance of commonly-observed Cloud Computing patterns-of-use to legal constraints and requirements.

Current European legislation

To clarify legal compliance in this field, we investigated commonly-observed Cloud use cases in collaboration with colleagues from the Tilburg University within the framework of the S-Cube project [2]. We considered the Data Protection Directive (DPD, 95/46/EC) of the European Union [1] – a commonly accepted and influential directive in the field of data processing legislation. This legislation is fundamental to Clouds as the consumer loses a degree of control over personal artefacts when they are submitted to the provider for storage and possible processing. To protect the consumer against misuse of their data by the provider, data processing legislation has been developed to ensure that the fundamental right to privacy is maintained. However, the distributed nature of Cloud Computing makes is difficult to analyse the data protection law of each country for common Cloud architecture evaluation criteria. The European DPD was designed to protect the privacy and all personal data collected for or about citizens of the EU, especially as it relates to processing, using or exchanging such data. The requirements of the DPD are expressed as two technology-neutral actors or roles that have certain responsibilities that must be carried out in order to fulfil the directive. These roles are naturally equivalent to service consumer and service provider roles found in distributed computing. The data controller is the natural or legal person which determines the means of processing of personal data, whilst a data processor is a natural or legal person which processes data on behalf of the controller. However, if the processing entity plays a role in determining the purposes or the means of processing, it is a controller rather than a processor.

Legal issues in Cloud use cases

We consider a generalized view of a Cloud Federation [3] that incorporates private, public, multi- and hybrid Cloud architectures derived from related Cloud standardization groups and European Cloud projects (NIST, ENISA, OPTIMIS). In this vision, interoperability is achieved by high-level brokering instead of bilateral resource renting, operated by a Service Provider (SP). In this situation different Infrastructure Providers (IP) may share or rent resources to provide a scalable infrastructure for SPs, which may be done transparently to the SPs. We have identified a series of use cases in federated Cloud architectures, in which legal issues may arise and necessary action should be taken in order to prevent violations. We found that the SP is mainly responsible for complying with data protection regulation. When personal data are transferred to multiple jurisdictions it is crucial to properly identify the controller since this role may change dynamically in specific actions. Information on the exact location of the processing establishments is also of great importance in these cases. Even if only one datacentre of a federation resides in the EU, the law of the appro-

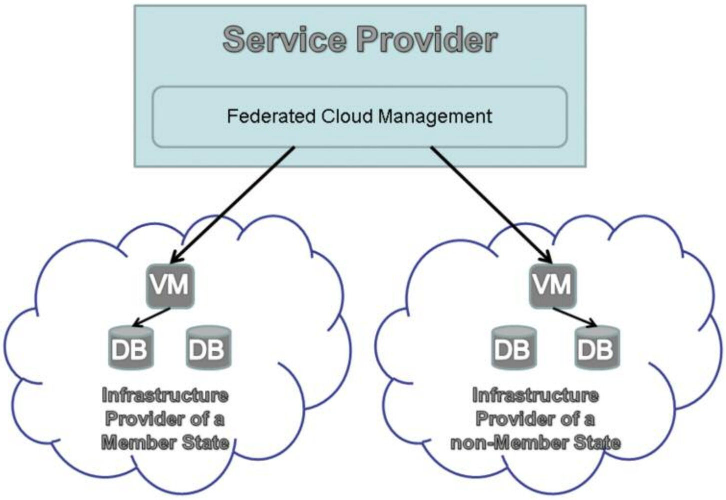

Service Provider
Federated Cloud Management
VM
DB
DB
Infrastructure Provider of a Member State
VM
DB
DB
Infrastructure Provider of a non-Member State

Figure 1: Data distribution regulations related to EU Member States with different infrastructure providers.

priate Member State (MS) of this datacentre must be applied by the SP. Figure 1 depicts a use case in which an SP provides a federated Cloud management in an MS. In this case different IPs are utilized, one of which is located in a non-MS. Since SP is the data controller and IPs are processors, the law of the SP's MS has to be applied, and the IP outside the EU has to provide at least the same level of protection as required by the national law of an MS. Otherwise, if a non-MS IP cannot ensure an adequate level of protection, the decision making process of SP should rule out this IP from provider selection.

Outlook

In summary, the currently effective European DPD is appropriate for determining the law applicable for data management in Cloud services when the data controller and processor roles are well identified. More problematic for companies, however, is process of applying the relevant law at a European scale, because the Member States have implemented the DPD in different ways. This issue has been recognized by the European Commission. They have proposed a reform of the European data protection rules in a regulation that will replace the current directive, the main goal being to strengthen the users' influence on their personal data.

Links:
http://www.lpds.sztaki.hu/CloudResearch
http://www.s-cube-network.eu/

References:

[1] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Official Journal L 281, pp. 31-50, Nov. 1995.

[2] Sz. Varadi, A. Kertesz, M. Parkin, The Necessity of Legally Compliant Data Management in European Cloud Architectures, Computer law and security review (CLSR), Elsevier (2012), doi:10.1016/j.clsr.2012.05.006

[3] A. Kertesz, A. Marosi, P. Kacsuk, Interoperable Resource Management for establishing Federated Clouds, In book: Achieving Federated and Self-Manageable Cloud Infrastructures: Theory and Practice, IGI Global (USA), pp. 18-35, 2012.

Please contact:

Attila Kertesz
SZTAKI, Hungary
E-mail: kertesz.attila@sztaki.mta.hu

Szilvia Varadi
University of Szeged, Hungary
E-mail: varadiszilvia@juris.u-szeged.hu

Providing Online Group Anonymity

by Oleg Chertov and Dan Tavrov

Protection of individual online privacy is currently a high profile issue. But, as important as it is, solving individual privacy issues does not eliminate other security breaches likely to occur if data aren't handled properly. Collective information about groups of people is also of vital importance and needs to be secured. Novel research into providing anonymity for particular groups highlights the necessity of privacy for groups.

The public is concerned about online privacy. Google's cloud computing services collect consumer data; search engine Yahoo keeps the data about Web searches; DVD-by-mail service Netflix keeps its users' movie records. Public awareness of these and similar cases force involved companies to improve their privacy policies.

Such a reaction is usually expected when it is necessary to preserve anonymity of a particular individual. But mining huge data collected online can violate privacy of a group of people as well. For instance, the data gathered on Facebook, if accessed and properly analysed (and this does occur), can violate privacy of a certain group of participants, even if this group is not explicit. An outstanding number of messages on a particular topic could point to a group of network members belonging to the same community, for example. This information can be valuable for tracking down criminal activities, for marketing, and so forth.

In other words, we face the problem of providing group anonymity of statistical data. Within our project conducted at the Applied Mathematics Department of the National Technical University of Ukraine “Kyiv Polytechnic Institute”, we have derived algorithms that provide group anonymity and have applied them to real statistical data. The basic features of our propositions are outlined below.

In general, providing group anonymity implies performing three steps. The dataset to be protected needs to be mapped onto appropriate data representation which can help define sensitive group data features. Then, such representation has to be appropriately modified in order to protect them. Afterwards, it is necessary to obtain modified dataset from the modified data representation.

Group anonymity heavily depends on data representation. The same dataset can be represented in many ways convenient for protecting the data against privacy breaches. Within our project, we proposed various ways of representing statistical data, each suitable for analysing particular data properties.

To explain this, let us consider a simple procedure of counting respondents possessing certain property and living in a certain area. For illustration, we can count the number of active duty military personnel living in different areas. A graph of the data reveals areas with denser populations of military personnel. The green line in Figure 1 represents such a graph, calculated for the military personnel distributed over statistical areas of the state of California, the US (according to the US Census 2000). Extreme values in the graph inevitably indicate the location of military bases. If this information is

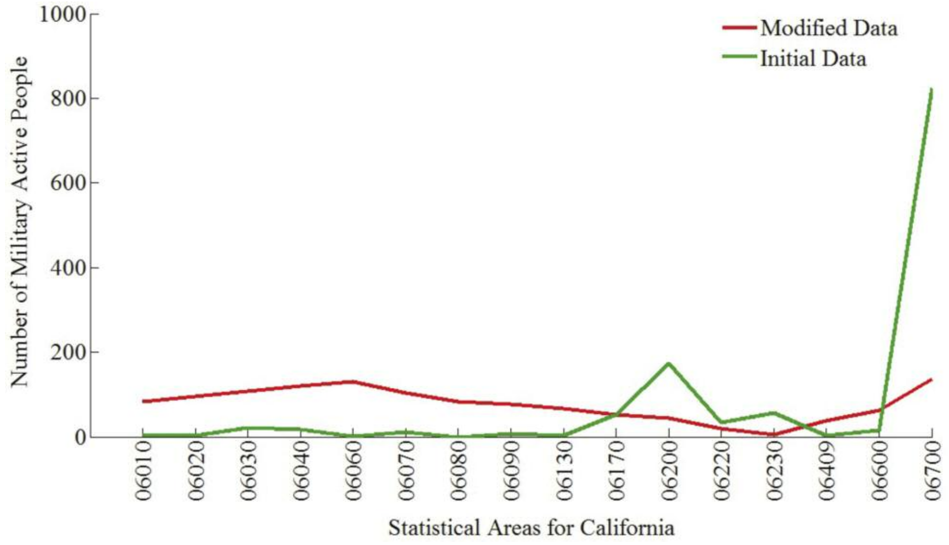


Figure 1: Numbers of active duty military personnel distributed by statistical areas of California. The green line shows the distribution computed using US population census 2000 microfile data provided by the US Census Bureau, http://www.census.gov/census2000/PUMS5.html. The red line shows distorted numbers obtained by solving the group anonymity problem using Daubechies wavelet of order 2.

required to be confidential, it is necessary to conceal the distribution.

One of the easiest ways to modify this data representation is to redistribute active duty military respondents by different areas to obtain a completely disseminate graph. Our research has demonstrated that an arbitrary distortion of the initial dataset leads to certain information loss which must be prevented.

We applied several methods each aiming to preserve a particular feature of the data. Wavelet analysis is a frequently used technique for splitting data into low- and high-frequency components. The low-frequency component, called “approximation”, stands for the smoothed version of the data, while the high-frequency components, referred to as “details”, tell much about intrinsic data properties. We propose to make full use of both approximation and details by playing around with the appearance of approximation which allows us to achieve the desired transformed distribution (Figure 1, red line). At the same time we leave details intact, thus guaranteeing preservation of the useful properties of the dataset.

After having obtained transformed distribution, it is necessary to accordingly modify the underlying dataset, introducing as less distortion as possible.

We intend to continue our research in the field of group anonymity to make better use of the information underlying the statistical data. In particular, we are considering applying fuzzy logic techniques to achieve anonymity, since statistical data consist mainly of various linguistic values which require proper processing.

We invite every interested ERCIM member to cooperate in conducting research in the group anonymity field which may be viewed as a logical succession of the previously completed Privacy Os-European Privacy Open Space project.

References:

- O. Chertov and D. Tavrov, “Group Anonymity,” Communications in Computer and Information Science, vol. 81, Part II, pp. 592–601, 2010.
- O. Chertov and D. Tavrov, “Data group anonymity: general approach,” International Journal of Computer Science and Information Security, vol. 8, № 7, pp. 1–8, 2010.
- O. Chertov, D. Tavrov, D. Pavlov, M. Alexandrova, and V. Malchikov, Group Methods of Data Processing, O. Chertov, Ed. Raleigh: Lulu.com, 2010, 155 p.

Please contact:

Oleg Chertov, Dan Tavrov
National Technical University of Ukraine "Kyiv Polytechnic Institute"
E-mail: chertov@i.ua, dan.tavrov@i.ua

How to Protect Data Privacy in Collaborative Network Security

by Martin Burkhart and Xenofontas Dimitropoulos

At ETH Zurich, we have developed the SEPIA library, which makes secure multiparty computation practical for many participants and large input data volumes. This paves the way for novel approaches in collaborative network security.

We currently have a fundamental imbalance in cybersecurity. While attackers are acting in an increasingly global and coordinated fashion, eg by using botnets, their counterparts trying to manage and defend networks are limited to examining local information. Collaboration across network boundaries would substantially strengthen network defense by enabling collaborative intrusion and anomaly detection.

Unfortunately, privacy concerns largely prevent collaboration in multi-domain cyberdefense. Data protection legislation makes data sharing illegal in certain cases, especially if personally identifying information (PII) is involved. Even if it were legal, sharing sensitive network internals might actually reduce security if the data should fall into the wrong hands. To address these privacy concerns, a large number of data anonymization techniques and tools have been developed. However, these anonymization techniques are generally not lossless. Therefore, organizations face a delicate privacy-utility tradeoff. While stronger sanitization improves data privacy it also severely impairs data utility.

As an alternative to simplistic data anonymization techniques, we started the SEPIA project [1, 2] at ETH Zurich a few years ago. SEPIA aims at providing practical secure multiparty computation (MPC) tools for collaborative network security. MPC is a cryptographic framework that allows running computations on data distributed among multiple parties, while provably preserving data privacy without relying on a trusted third party. In theory, any computable function on a distributed data set is also securely computable using MPC techniques. Unlike anonymization, MPC gives information-theoretic guarantees for input data privacy. A well-known example of MPC is the Millionaire’s Problem: Two millionaires want to know which of them is richer without revealing their fortunes to each other. Using MPC, the millionaires can compute the answer to their question, without learning anything else.

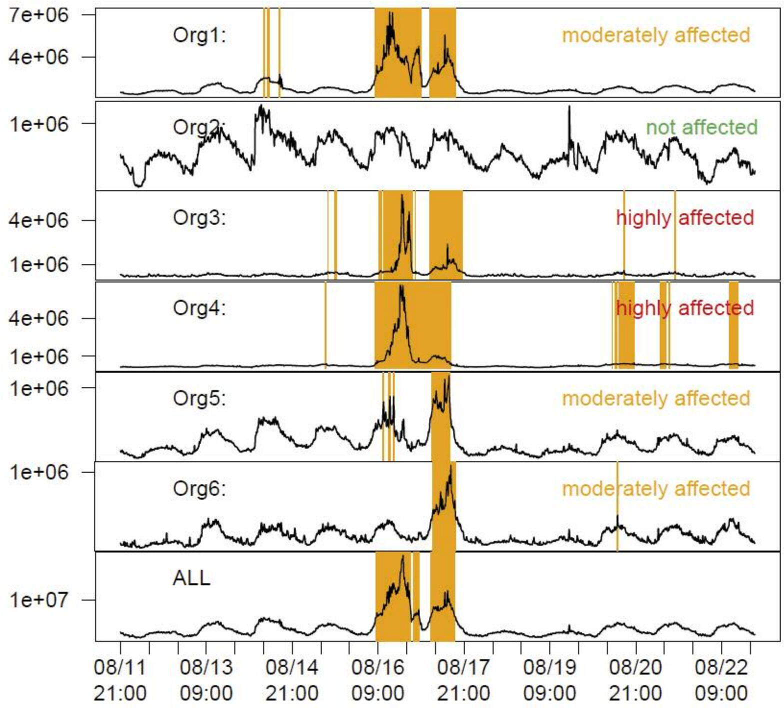


Figure 1: Visibility of the Skype 2007 anomaly in counts of Netflow records across six networks. Collaborative protocols aid in identifying root causes by combining local views into a bigger picture.

Although MPC has been studied substantially for almost 30 years, building solutions that are practical in terms of computation and communication cost is still a major challenge, especially if input data are voluminous as in our scenarios. Therefore, we developed new MPC operations for processing high volume data in near real-time. The prevalent paradigm for designing MPC protocols is to minimize the number of synchronization rounds. However, the resulting protocols tend to be inefficient for large numbers of parallel operations, requiring hours or even days to compute simple functions. By optimizing the design of basic operations, we managed to significantly reduce the CPU time and bandwidth consumption of parallel MPC operations. SEPIA's operations are between 35 and several hundred times faster than those of comparable MPC frameworks.

SEPIA provides a layer of basic MPC operations for addition, multiplication, and different types of comparisons that can be used through an API to build more advanced composite protocols. On top of the basic operations, it provides a layer of composite protocols designed for network security and monitoring applications that require the aggregation of events and statistics from multiple domains. For example, the top-k composite protocol aggregates lists of key-value pairs from multiple parties and reveals only the keys with the k largest aggregate values. It can be used to implement an efficient IDS alert correlation mechanism to support multi-party collaboration, for instance in the fight against botnets, without requiring a trusted third party. So far SEPIA has implementations of ten such composite protocols including protocols for private set operations and distributed network anomaly detection.

To learn how SEPIA can be useful in practice, we applied it to traffic data from 17 customer networks of the Swiss NREN (SWITCH) collected during the global Skype outage in August 2007. Figure 1 shows the Netflow count across six networks and the aggregate count as computed by SEPIA at the bottom (ALL). The Skype anomaly is clearly visible in the middle. By comparing their local view of the anomaly with the aggregate view, the organizations can quickly decide whether the scope is local or distributed. Also, they can assess how much the local network is affected compared to others and sometimes even profit from early warnings (eg organization 6 was hit one day later than the others). Due to their privacy-awareness, MPC protocols allow the correlation of additional sensitive features, such as the top ports, IP addresses, or IDS alerts during normal and anomalous periods. Such aggregate information beyond the limited local view enables operators to quickly get the big picture of what is going on and act accordingly.

These results are indeed promising and we believe that MPC will play an important role in the future of collaborative network security. SEPIA is already being used by other researchers, for example for practical PIR in electronic commerce [3]. In addition, the Swiss Commission for Technology and Innovation and IBM Research are funding an on-going project that aims at integrating SEPIA with a network monitoring commercial system from IBM thereby enabling privacy-preserving multi-domain flow analysis, which could lead to one of the first commercial products that use MPC.

Link:
[1] http://www.sepia.ee.ethz.ch/

References:

[2] M. Burkhart and X. Dimitropoulos. Privacy-Preserving Distributed Network Troubleshooting - Bridging the Gap between Theory and Practice. ACM Transactions on Information and System Security (TISSEC), 14(4), Dec. 2011.

[3] R. Henry, F. Olumofin, and I. Goldberg. Practical PIR for electronic commerce. ACM CCS, 2011.

Please contact:

Martin Burkhart
Xenofontas Dimitropoulos
ETH Zurich
E-mail: burkhart@tik.ee.ethz.ch, fontas@tik.ee.ethz.ch
Tel: +41 44 632 70 04

Personal Data Server: Keeping Sensitive Data under the Individual's Control

by Nicolas Anciaux, Jean-Marc Petit, Philippe Pucheral and Karine Zeitouni

In the IT world every piece of information is just "one-click away". This convenience comes at a high indirect price: the loss of the user's control over her personal data. We propose a simple yet effective approach, called Personal Data Server, to help protect the user's data.

An increasing amount of personal data is gathered on servers by administrations, hospitals, insurance companies, etc. Smart devices all around us also produce transparently spatio-temporal sensitive information (eg healthcare monitoring, smart buildings, road pricing). In the meantime, more and more digitized data is delivered to the user (salary forms, invoices, phone call sheets, banking statements, etc). While primary copies of these data are kept by the issuer information systems, citizens themselves often rely on Internet companies to reliably store secondary copies and make them available online. Unfortunately, there are many examples of privacy violations arising from negligence, abuse and attacks and even the most secured servers are not spared.

We draw a radically different, highly decentralized, vision of the management of personal data. It builds upon the emergence of new devices known as Secure Tokens combining the security of smart cards and the storage capacity of NAND Flash chips (eg mass storage SIM cards, secure USB sticks, smart sensors). This unprecedented conjunction of portability, security and mass storage holds the promise of a real breakthrough in the management of personal data.

The idea is to embed in Secure Tokens, software components capable of acquiring, storing and managing securely personal data. This forms a full-fledged Personal Data Server (PDS) remaining under holder's control. PDS is not a simple secure repository of personal data. It must allow the development of powerful, user-centric and privacy-preserving applications thus requiring a well organized and queryable representation of user's data. It must also provide the data holder with a friendly control over the sharing conditions related to her data. PDSs must finally provide traditional database services like durability and query facilities and must be able to interoperate with external data sources in a secure manner.

With appropriate infrastructure, PDSs enable the vision depicted in Figure 1. John's personal data, delivered and certified by different sources, are sent to his PDS which can then serve data requests from various applications. While protecting personal information on the issuer side will remain an open problem, the PDS vision enables executing all these applications under the full control of individuals.

Private applications are run by the holder herself and inherit her privileges

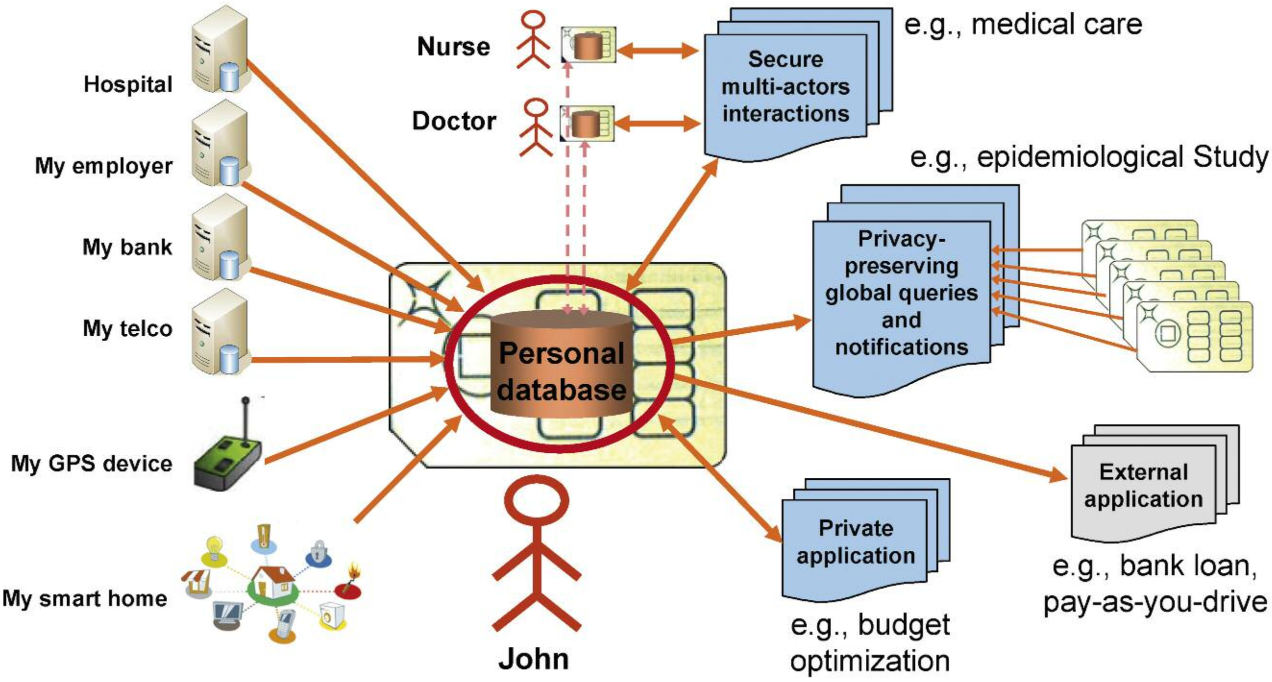

Hospital
My employer
My bank
My telco
My GPS device
My smart home
Nurse
Doctor
Personal database
John
Secure multi-actors interactions
e.g., medical care
Privacy-preserving global queries and notifications
e.g., epidemiological Study
External application
e.g., bank loan, pay-as-you-drive
Private application
e.g., budget optimization

Figure 1: Personal Data Server Architecture

(eg a budget optimizer). External applications/services declare collection rules specifying which data is required by which business rules (eg salary and tax information for a bank loan, aggregates of GPS data for a PayAsYouDrive system). The PDS then computes the minimum amount of data to be disclosed and shows it to the holder who can finally opt-in/out for this service. Conversely, data provenance is certified by the PDS and can be checked by the service. Privacy-preserving global computations are large-scale treatments made on a population of PDSs (eg aggregate queries or anonymized release for an epidemiological study), providing privacy guarantees (eg differential privacy, k-anonymity) enforced by all PDSs in a secure multi-party way. Finally, collaborative applications can exchange personal data among PDSs (eg patient's data exchanged among doctors, personal files exchanged among a community of friends or colleagues). Sharing situations can be bounded in time thanks to data retention rules and can be audited afterwards. Secure collaborative scenarios can be achieved because all participant PDSs are tamper-resistant and trusted.

Converting the PDS vision into reality introduces three main scientific challenges: (1) new database techniques must be devised to efficiently manage embedded data while tackling the strong hardware constraints inherent to Secure Tokens ; (2) a rich and intuitive model must be provided to help individuals protect all facets of their privacy and proof of legitimacy must be provided for any data entering/leaving a PDS; (3) the traditional functions of a central server must be re-established in an atypical environment combining a large number of highly secure but low power Secure Tokens with a powerful but unsecured infrastructure.

Our hope is that the PDS approach will provide a credible alternative to the systematic centralization of personal data on servers and will pave the way for new privacy-by-design architectures.

This work is supported by the KISS ANR project which groups academics (INRIA, LIRIS, UVSQ), industry (Gemalto, CryptoExperts) and administrations (Yvelines District General Council-CG78). A platform prefiguring the PDS vision is being experimented in the field under grant DMSP. It implements a secure and portable patient's folder improving social and medical care coordination for elderly people.

Links:
KISS project: http://blog.inria.fr/kiss/
DMSP project: http://www-smis.inria.fr/_DMSP/accueil.php

Reference:
T. Allard et al, 'Secure Personal Data Servers: a Vision Paper', 36th International Conference on Very Large Data Bases, PVLDB 3(1): 25-35, 2010.

Please contact:
Philippe Pucheral (KISS project)
University of Versailles & Inria ,
France
E-mail: Philippe.Pucheral@inria.fr

Nicolas Anciaux (DMSP project)
Inria, France
E-mail: Nicolas.Anciaux@inria.fr

The Minimum Exposure Project: Limiting Data Collection in Online Forms

by Nicolas Anciaux, Benjamin Nguyen and Michalis Vazirgiannis

When requesting bank loans, social care, tax reduction, and many other services, individuals are required to fill in application forms with hundreds of data items. It is possible, however, to drastically reduce the set of completed fields without impacting the final decision. The Minimum Exposure Project investigates this issue. It aims at proposing an analysis, framework and implementation of an important privacy principle, called Limited Data Collection.

Personal data collection is a prerequisite to well-tailored services, which are in the interest of both service provider and applicant. A classical way to collect such data is to issue application forms. When considering privacy from the applicant's point of view it is unquestionable that the personal information harvested in these forms must be reduced to a minimum necessary to make the correct decision.

Minimizing the data collected has also become a financial issue for service providers. Collected records are threatened by data breaches, which are not a marginal problem. In 2011, various sources, such as the Open Security Foundation, have reported tens of millions of personal records subject to such breaches. The average cost per exposed record has been estimated at $194 by security organizations such as the Ponemon Institute. Moreover, recent laws enacted worldwide (in 46 US states and EU) now compel companies to publicly report incidents and assist victims in minimizing their effect.

While better securing servers is a hot research topic, it is still not possible to provide a 100% secure environment. Our project proposes an orthogonal approach. With regards to legislation, our work assumes a strict understanding of the privacy principle termed “Limited Data Collection”, which states that requested sets of personal data must be limited to the minimum necessary to achieve the purpose the user consents to.

The precise contribution of the Minimum Exposure project is to provide guidelines, algorithms, and a framework to implement Limited Data Collection, since nothing exists currently.

Current practices fail to comply with this principle for the following reason: It is impossible to distinguish a priori which data will be useful (or not) to make the decision at the time the application form is filled; not only does information harvesting depend on the purpose, it also depends on the contents. Consider the (simple) collection rules based on the following tax rate reduction example: revealing an income under *$30,000* and an age below *25* may be enough, or simply an income below *$10,000*, regardless of age. Alternatively, revealing simply a sufficient number of dependants (eg two) could suffice. For a user with values u_1=[*income=$25,000, age=21, nb_dependants=1*] the minimum data set would be [*income, age*]. For a user with u_2=[*income=$40,000, age=35, nb_dependants=2*] it would be [*nb_dependants*]. Hence, the organization issuing the form cannot specify a priori a minimum set of attributes needed to make its decision since this decision depends on looking at the value of all attributes available.

To circumvent this issue, our framework proposes to bind collection rules with application forms. The framework is depicted below, and illustrated by the following scenario : When a user wants to apply to a service, she ① downloads the application forms provided by the service provider and fills in the form given the information she owns, ② runs a Minimum Exposure process to compute the minimum set of data items to provide in the application form to obtain the benefits she wants using the collection rules, ③ validates and signs the resulting application form and sends it to the service provider. The service provider can ④ run its decision processes using the contents of the form and calibrate its offer.

The Minimum Exposure project faces challenging problems at the intersection

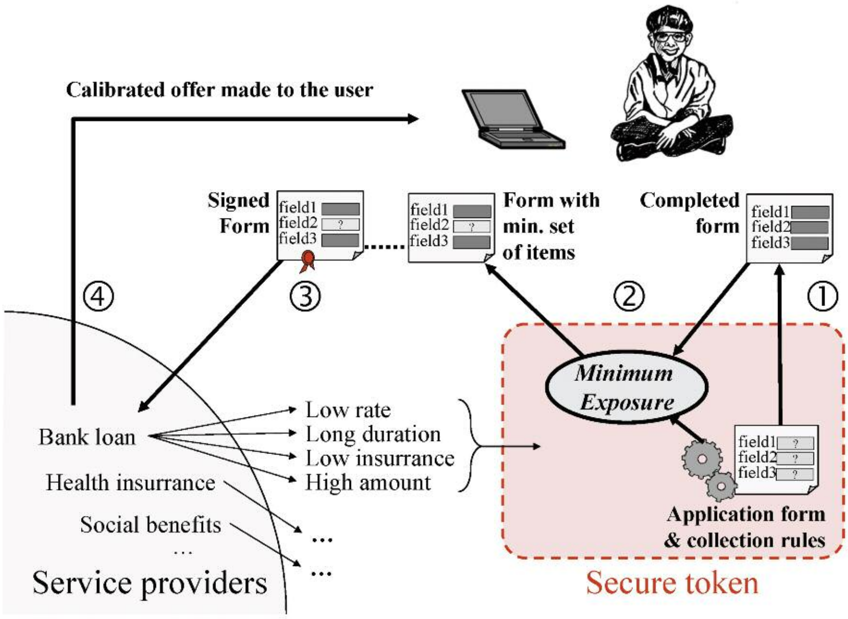


Figure 1: Scenario for limiting data collection

of data mining, secure data computation, and operational research. First, the collection rules attached to application forms must cover any decision making system ranging from simple disjunction of conjunction of predicates enacted by law (eg e-administration scenario) to highly complex systems based on data mining techniques (eg neural networks used in credit scoring for bank loans applications). Second, the decision making process is often private for the service provider, and the related collection rules must not be revealed, leading to adoption of secure tokens (like smart cards) in the architecture. Third, identifying the minimum set of information to be sent while still preserving the final decision is a NP-Hard problem. Application forms can be very large in practice (eg hundreds fields in loan application forms or tax declarations), leading to introduction of approximation algorithms based on heuristics adapted to the topology of these rules.

The Minimum Exposure project is a collaboration started in 2011 between INRIA, University of Versailles and Ecole Polytechnique. It is partially funded by the DIGITEO Letevone Grant and by the INRIA CAPPRIS initiative.

Link: http://project.inria.fr/minexp/

Reference:

[1] N. Anciaux, B. Nguyen, M. Vazirgiannis, “Limiting Data Collection in Application Forms: A real-case application of a Founding Privacy Principle”, in IEEE 10th annual conference on Privacy, Security and Trust (PST), 2012, to appear.

Please contact:

Benjamin Nguyen
University of Versailles and INRIA, France
E-mail: benjamin.nguyen@inria.fr

Linking a Social Identity to an IP Address

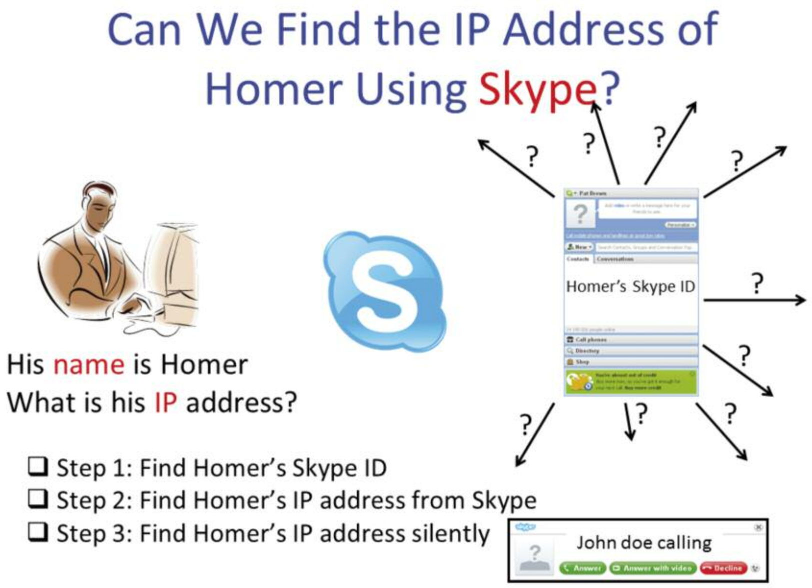
by Arnaud Legout and Walid Dabbous

Linking a social identity such as a name to an IP address is generally believed to be difficult for an individual with no dedicated infrastructure or privileged information. Although an individual’s ISP has access to this information, it is kept private except in the case of a legal decision. Similarly, some big Internet companies such as Facebook and Google might be privy to this information but it will never be communicated as it is an industrial secret used for targeted advertisements. In the context of the bluebear project, we show that it is possible for an individual to inconspicuously make the link between social identity and IP address for all Skype users.

The privacy threat that exists online is a growing concern. Most academics and journalists focus on the threat posed by the huge amount of data collected by big companies such as Google or Facebook. However, the case of individuals, with no dedicated infrastructure or privileged information, trying to infringe Internet users’ privacy is largely overlooked owing mainly to the misconception that it is impossible for a single individual to spy on Internet users at a large scale. The goal of the bluebear project, led by researchers at Inria in collaboration with researchers at the Polytechnic Institute of New York University (NYU-Poly), is to explore whether individuals can infringe Internet users privacy at a large scale.

In a first study, we have shown that an individual can monitor all BitTorrent downloads in real time for all Internet users without any dedicated infrastructure. To prove this, we collected 148 million IP addresses downloading more than one million contents for 103 days, and managed to identify 70% of users who first insert content in BitTorrent. We have also shown that using Tor, the anonymizing overlay, does not help; Tor does not protect against the exploitation of an insecure application (eg BitTorrent) to reveal the IP address of a TCP stream. Even worse, because Tor sends application data together over a single circuit, the IP address found for a given application can be associated with all other applications of the same circuit.

Although profiling Internet users by IP addresses raises ethical concerns, it pales in comparison to profiling by social identity. Associating a social identity to a profile enables very severe privacy infringements that might lead to real world attacks such as blackmail or targeted phishing attacks. We used Skype as a case study to show that an individual can link a social identity and an IP address. In particular, we have shown that an individual can inconspic-



Can We Find the IP Address of Homer Using Skype?

His name is Homer
What is his IP address?

❑ Step 1: Find Homer’s Skype ID
❑ Step 2: Find Homer’s IP address from Skype
❑ Step 3: Find Homer’s IP address silently

Figure 1: High level description of the social identity and IP address linkage exploiting Skype.

uously retrieve the IP address of any Skype user. As 88% of Skype users provide a name associated with the account and 82% provide additional information like a country or Web page, most Skype accounts can be directly associated with a social identity, thus making it possible to link social identity and IP address. In addition, we can follow the mobility of Skype users. Interestingly, on a sample of 10,000 random Skype users, we observed that over a two week period 4% of them moved from one country to another, 19% moved from one ISP to another, and 40% moved from one city to another. Therefore, we identified real mobility patterns for a regular Skype user.

Even more of a concern, by tracking the mobility of an Internet user along with the mobility of acquaintances (retrieved from a social network profile like Facebook or LinkedIn) it is possible to track social interactions, that is who an Internet user meets and where. This has worrying implications for both the personal and professional lives of Internet users.

The technique we use to map a social identity to an IP address is based on the identification of specific communication patterns that are inherent to any peer-to-peer system. Therefore, we do not rely on a specific bug in Skype that exposes the IP address, but on the intrinsic peer-to-peer architecture used by Skype and the lack of privacy of the IP protocol used on the Internet. As a consequence, the kind of attacks we have documented can be adapted to most systems that use a peer-to-peer architecture. In addition, it is extremely hard, without a major architectural change to Internet communications, to make such attacks impossible.

Our goal in the next few years is to participate in the design of a more secure, privacy preserving Internet communication infrastructure.

Link:
http://planete.inria.fr/bluebear

References:

[1] S. Le Blond, C. Zhang, A. Legout, K. Ross, and W. Dabbous, "I Know Where You are and What You are Sharing: Exploiting P2P Communications to Invade Users' Privacy". In Proc. of ACM SIGCOMM/USENIX IMC'11, Nov. 2--3, 2011, Berlin, Germany. http://hal.inria.fr/inria-00632780/en/

[2] S. Le Blond, A. Legout, F. Lefessant, W. Dabbous, M. Ali Kaafar, "Spying the World from your Laptop - Identifying and Profiling Content Providers and Big Downloaders in BitTorrent". In Proc. of LEET'10, April 27, 2010, San Jose, CA, USA. http://hal.inria.fr/inria-00470324/en/

Please contact:

Arnaud Legout; Inria, France
E-mail: arnaud.legout@inria.fr

Reidentification for Measuring Disclosure Risk

by Vicenç Torra and Klara Stokes

The release of confidential data to third parties requires a detailed risk analysis. Identity disclosure occurs when a record is linked to the person or, more generally, the entity that has supplied the information in the record. A re-identification method (eg record linkage) is a tool which, given two files, links the records that correspond to the same entity. In the project Consolider-ARES we study re-identification methods, their formalization and their use for measuring disclosure risk.

When a database is released to third parties, the leakage of confidential information about the individuals in the database is an important issue. One common approach for solving this problem is to mask the data in order to destroy possible links between the confidential information and the individuals, at the same time preserving some utility of the data. In this process the anonymity of the individuals should be maximized and the information loss simultaneously minimized. In order to define good masking methods, it is essential to quantify anonymity and information loss. We believe that in most situations anonymity of individuals should be evaluated and ensured before information loss is taken into account, otherwise the anonymity is likely to suffer, as has been illustrated on several occasions. Consequently, our research focuses on quantifying and understanding anonymity in data privacy.

Identity disclosure occurs when a record in the database is linked to the person or organization whose data is in the record. In data privacy, re-identification methods (for example record linkage) are used to evaluate disclosure risk, while data protection methods are developed to avoid, or reduce the chance of, identity disclosure. The idea of re-identification is pervasive in data privacy, and some concepts as k-anonymity can be understood in the light of re-identification methods and record linkage.

Record linkage focuses on the case of two databases with information about the same individuals, and linkage is done at the record level. Re-identification is a more general term which encompasses the linkage of other objects as attributes and includes schema matching for example.

In the Consolider-ARES project we study re-identification methods. We address issues ranging from their formalization to algorithms to make re-identification effective. Our study includes the following topics:

1. Formalization of re-identification algorithms. This approach tries to answer the question of which algorithms are correct re-identification algorithms. Our formalization is based on imprecise probabilities and compatible belief functions. Only algorithms that return probabilities compatible with a true probability can be properly called re-identification algorithms. This construction permits us to revise probabilistic record linkage.
2. Optimal re-identification algorithms. Given two databases consisting of

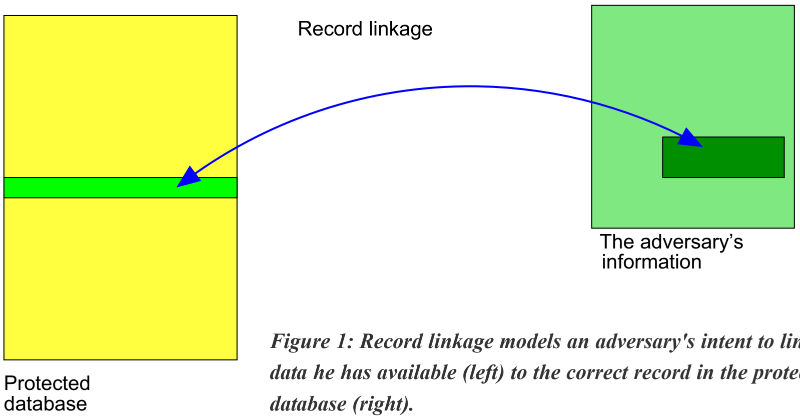


Figure 1: Record linkage models an adversary's intent to link the data he has available (left) to the correct record in the protected database (right).

information about the same individuals, when the correct links between the two databases are known, we can formalize the optimal re-identification problem in terms of the number of incorrect links. That is, given a parameterization of the algorithm, we define the objective function as the number of incorrect links. Then the optimal re-identification algorithm is the one that minimizes this objective function – the number of incorrect links. Any solution of this optimization problem can be used as a measure of risk of the worst-case scenario. Machine learning tools and optimization techniques can be used to find this optimal solution.

3. K-confusion. We say that a data protection algorithm satisfies k-confusion if it causes any re-identification method to return at least k possible candidates. This is a generalization of k-anonymity. In k-anonymity it is required that there are at least k records that are equal over the attributes that are assumed to be useful for the adversary's re-identification process. As k-confusion is a generalization of k-anonymity, it allows for more flexibility and can therefore be used in order to achieve masked data with lower information loss, without yielding on the degree of anonymity that is provided.

Within the Consolider-ARES project, re-identification methods have been applied to different types of databases. These databases include standard numerical and categorical tables, but also databases with time series, series for locations, and graphs to represent online social networks. Re-identification algorithms have been used to measure the disclosure risk of the outcome of data protection methods.

Partial support by the Spanish MEC (projects ARES – CONSOLIDER INGENIO 2010 CSD2007-00004 – and RIPUP– TIN2009-11689) is acknowledged.

References:

[1] V. Torra, “Privacy in Data Mining”, in Data Mining and Knowledge Discovery Handbook, O. Maimon, L. Rokach, Eds., 2nd Edition, Springer, 2010, pp. 687-716.
DOI:10.1007/978-0-387-09823-4_35

[2] D. Abril, G. Navarro-Arribas, V. Torra, “Improving record linkage with supervised learning for disclosure risk assessment”, Information Fusion, Volume 13, Issue 4, October 2012, Pages 274–284.
DOI:10.1016/j.inffus.2011.05.001

[3] K. Stokes, V. Torra, “Reidentification and k-anonymity: a model for disclosure risk in graphs”, Soft Computing, in press.
DOI:10.1007/s00500-012-0850-4

Please contact:

Vicenç Torra
IIIA-CSIC, Spain
Tel: +34 935809570
E-mail: vtorra@iiia.csic.es

Klara Stokes
Universitat Oberta de Catalunya, Spain
Tel: +34 93 450 54 17
E-mail: kstokes@uoc.edu

The Real Value of Private Information – Two Experimental Studies

by Marek Kumpošt and Vashek Matyáš

As a part of our research on privacy protection and identity management, we conducted two experiments to find out how people value their private information. Privacy and control of private information sharing/flow is becoming a crucial part of everyday “online” life. But still, people seem to be prepared to disclose private data for a very modest reward – loyalty cards, for example, allow profiling of customer behaviour and use of this information (to create, e.g., personalized advertisements. Search engines and social networks can track users’ browsing activities via embedded sharing buttons. This is a very common technique and even if we are not a member of a social network, there is a lot of evidence about our browsing history. This can also be used for providing customized content.

Privacy protection should be an important aspect of everyone’s real life, but what is the “real” value of private information? This question was our motivation for organizing two experiments. Both projects were undertaken within the Future of Identity in the Information Society (FIDIS) Network of Excellence activities. In both experiments, we used a cover story to hide the real purpose of a questionnaire. This approach enabled us to avoid the

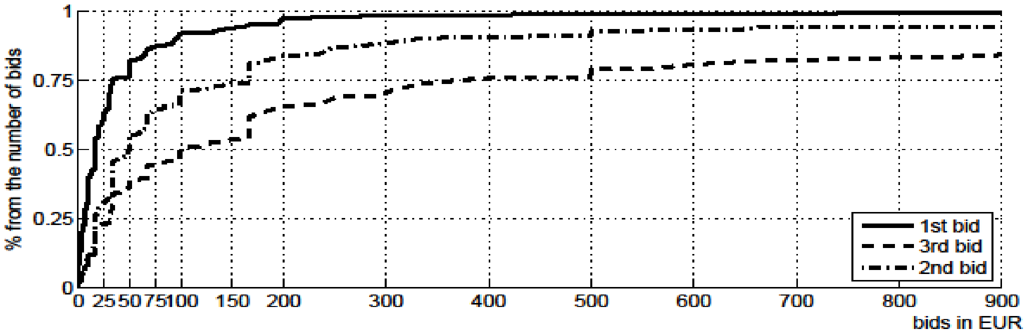


Figure 1: Distribution of bids for all three study scenarios

difficulties associated with obtaining realistic data when people are asked directly about the value of their private information. Details of the experimental design can be found in [1, 2].

In the first experiment [1], we questioned a sample of 1,200 people from five EU countries about taking a part in a (fictitious) study of mobile phone usage. Each participant could place a bid; a financial reward he/she wanted to receive from the study. The truth was, that we were interested in the monetary value people would attach to a month of location data (collected through their mobile phones every five minutes). The structure of all questionnaires can be found in [1]. We extracted valuations for different kinds of data usage: qualitative – academic or commercial purposes; and quantitative – one month and one year.

Figure 1 shows bid distributions of people who provided data for all three scenarios (1 – academic usage; 2 – commercial usage, one month; 3 – commercial usage, one year). The x-axis shows the value of bids in euros and the y-axis shows the fraction of bidders who entered bids of value lower or equal to a given amount. We standardized the data to account for the buying power in individual countries. The median of bids increased approximately twofold when the data were to be used for commercial purposes. The extension from one month to one year resulted in another twofold increase. This indicates that participants were more sensitive to the purpose of the data collection than the duration and quantity of data collected. The median bid for the first scenario (academic use of data) was €43. More detailed results can be found in [1].

Our second experiment [2] followed the same design but with the different private information being priced. The main goal of the study was to find out how people value information about their use of online communication tools such as email or instant messaging. The real purpose of the study was concealed by a cover story stating that it was to be a two-week sociological study about the use of online communication (using a special tool installed on a participant's machine). Monitoring scenarios were: email traffic data (no message body); instant messaging traffic data; and all traffic data. Data would then be used for academic purposes – line 1; commercial purposes – line 2; or governmental purposes – line 3 in Figure 2.

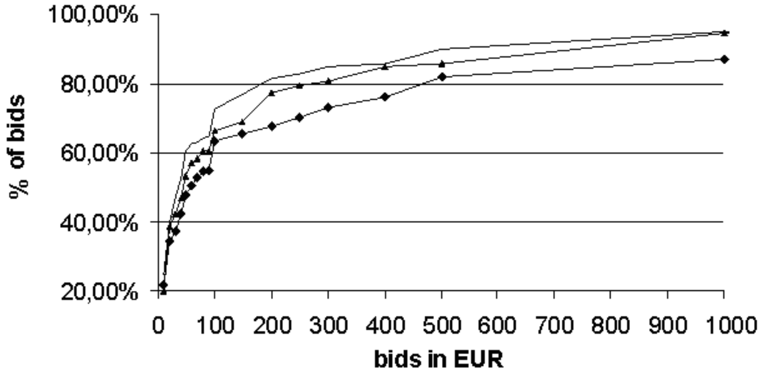


Figure 2: Distribution of bids for all communication data all three study scenarios

Figure 2 shows the distribution of bids for all tracking data. Participants did not differentiate between the type (email versus all) of tracking data and the median monetary compensation for email traffic data was €30 (the same price for instant messaging). All tracking data are "more expensive", the median being €50. More information and comprehensive analyses can be found in [2]. Interestingly, the participants (young generation) were less willing to share the same data for commercial purposes than for academic use and even less so for governmental efforts to improve terrorist activity tracking tools. This was shown by both declining number of participants willing to provide their data for such purposes, as well as by rising valuation of the data by those who would still be willing to participate.

Link:
http://www.fidis.net

References:

[1] D. Cvrček, M. Kumpošt, V. Matyáš, G. Danezis, "A Study on The Value of Location Privacy". In: Fifth ACM Workshop on Privacy in the Electronic Society. USA: ACM, 2006. pp. 109-118. ISBN 1-59593-556-8.

[2] V. Matyáš, M. Kumpošt, "Monetary valuation of people's private information". In: Privacy and Usability Methods Pow-wow (PUMP) 2010; 2010. 7 pp. 2010, University of Abertay Dundee.

Please contact:

Vashek Matyáš
Masaryk University, Brno – CRCIM, Czech Republic
Tel: +420 549 49 5165
E-mail: matyas@fi.muni.cz

Microarrays - Innovative Standards in a Changing World: the Case for Cloud

by Jane Kernan and Heather J. Ruskin

Large public Microarray databases which store extensive data on different genes, are still in high demand. Integration of the different data types generated poses many problems, not least in terms of quality assessment, scope and interpretation. Old and new paradigms for management co-exist within the field, including Data Warehouses, Database Clusters, Federations and Cloud computing, which need to be sensitive to both historical provision and future directions. So, what is involved?

Microarrays and other high throughput technologies provide the means to measure and analyse expression levels of a large number of genes and samples. In early stage development of microarray technologies, limitations in the initial structure of the datasets rapidly became evident, together with the demand for flexible computational tools and visualisation capability to extract meaningful information. Requirements for shared data access were also recognised. These initial demands have led to new developments in data storage, analysis and visualisation tools, and have facilitated identification of previously undetected patterns in complex datasets. Solving one set of problems is rarely an end in itself, however, and so it has proved in this case, with new data types (typically from Next Generating sequencing methods), now on the increase, the problem is one of reconciling/integrating the different types to maximum advantage.

In any consideration of data quality in the current context, recognition of the major contribution to the success and wide use of microarray technology is due to MIAME, (Minimum Information About a Microarray Experiment), a set of standards for microarray experiment annotation, launched in 1999 by the Microarray Gene Expression Data (MGED) group, (now known as Functional Genomics Data Society (FGED). The MGED Ontology working group are continuing to develop a common set of terminologies, the microarray ontology (MO) and Gene Ontology (GO) to facilitate automated querying and exchange of microarray data. Furthermore, the Microarray Quality Control (MAQC) Consortium have developed quality control standards that operate to ensure the reliability of experimental datasets.

Microarray technology has also evolved from the flat file systems and spreadsheets of the early 1990's to the various database management systems, both public and private, that are accessible through networks and the Internet, sharing data in a more structured and reliable format. With such a variety of web resources, databases, data models, and interfaces, information gathering on properties for a specified gene list is non-trivial as integration is needed to access data, located at numerous remote sources, using different interfaces and query results, which are displayed in various formats. The ideal, clearly, is integration of multiple information sources and a simplified interface, but there is no single solution, to date, which allows comprehensive data on a specific gene to be collated. One federated system, overseen by the National Cancer Institute Center for Biomedical Informatics and Information Technology (NCI-CBIIT) is CaBIG (the Cancer Biomedical Informatics Grid http://cabig.nci.nih.gov/), where a suite of tools provides for simplification of data input and sharing across the Cancer Biomedical Informatics Grid, using a federated model of local installations. A Cloud hosting environment for health

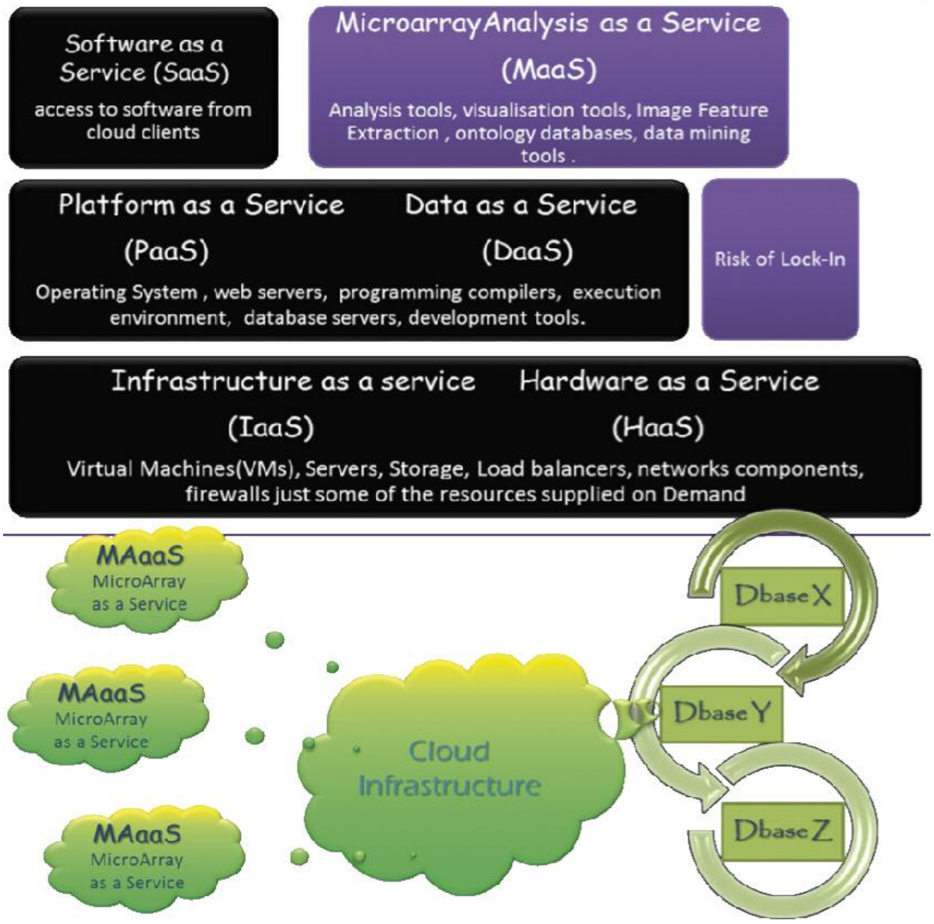

Software as a Service (SaaS)
access to software from cloud clients
MicroarrayAnalysis as a Service (MaaS)
Analysis tools, visualisation tools, Image Feature Extraction , ontology databases, data mining tools .
Platform as a Service (PaaS)
Data as a Service (DaaS)
Operating System , web servers, programming compilers, execution environment, database servers, development tools.
Risk of Lock-In
Infrastructure as a service (IaaS)
Hardware as a Service (HaaS)
Virtual Machines(VMs), Servers, Storage, Load balancers, networks components, firewalls just some of the resources supplied on Demand
MAaaS MicroArray as a Service
MAaaS MicroArray as a Service
MAaaS MicroArray as a Service
Cloud Infrastructure
Dbase X
Dbase Y
Dbase Z

Figure 1 Possible Architecture for Microarray Cloud Computing

care data, (attracting a monthly subscription), is also provided.

Is cloud computing the answer to integration? Does cloud computing make economic sense? Its selling point is that using a cloud computing facility means you do not have to pay for distribution, infrastructure installation or maintenance costs. You only pay for what you use. Software updates and compatibilities are no longer an issue and this is achieved by use of Virtual Machines (VMs). VMs provide the ability for resource pooling. One cloud computing service provider is Amazon (Elastic Cloud Computing EC2) offers a variety of bioinformatics-oriented virtual machine images as well as providing several large genomic datasets in its cloud (copies of GeneBank DB and Genome databases from Ensembl), once portability and interoperability standards are introduced this may be a viable option.

When choosing to migrate to a cloud one of the main concerns is that important data will reside on a third parties server, other concerns are, security and lack of control over the system, the trust-worthiness of the service provider, and the lack of portability standards, which may heighten the risk of company lock-in for customers wishing to move vendors. How simple a process is moving vendors going to be? Governing authorities such as the IEEE Standards Association (IEEE-SA) have formed two new Working Groups (WGs) IEEE P2301 (Cloud Portability and Interoperability) and IEEE P2302 (Standard for Inter cloud Interoperability and Federation) to try and standardise the approaches of critical areas such as portability, interoperability interfaces and file formats.

Migrating to cloud computing is not a trivial task, so the question is can Cloud with its combination of Computational Power, Big Data and Distributed Storage, Horizontal Scalability and Reliability be harnessed to provide a one-stop-shop for Microarray Analysis? Can complex multi-faceted searches on a distribution of connected databases be supplied on demand to microarray researchers through a Microarray Analysis-as-a-Service type client process (which is a "one-to-many" model whereby an application is shared across multiple clients). This research is on-going and future work will report on interoperability and migration issues of these biological databases when combined with the latest cloud technologies.

Links/References:

http://opencloudmanifesto.org/Cloud_Computing_Use_Cases_Whitepaper-4_0.pdf
http://www.biomedcentral.com/1471-2105/12/356
CaBIG: http://cabig.nci.nih.gov/

Please contact:

Jane Kernan
CloudCORE Research Centre
Dublin City University, Ireland
E-mail: Jane.Kernan@computing.dcu.ie

Tackling IT Challenges in Multi-Participant Hub-and-Spoke Logistics Networks

by Zsolt Kemény and Elisabeth Ilie-Zudor

In the past, fast and affordable shipping of small consignments was perceived by the logistics industry as a conflict of requirements. Recent decades, however, have shown that bundling and rebundling of shipments enables certain types of logistics networks to comply with both criteria at the same time. Usually, these networks follow a multi-level hub-and-spoke layout with a high-throughput central player and independent and agile local collection and delivery partners. Meanwhile, however, performance limits have surfaced, most of them owing to organisational heterogeneity and resulting degradation of process transparency. Since late 2010, the FP7 EU project ADVANCE has been tackling these challenges and providing a re-usable framework for modelling and solving problems inherent to hub-and-spoke networks.

Past decades have demonstrated the feasibility of networks with an efficient core of one or more hubs, and a periphery of flexible local pick-up and delivery services acting as spokes. Inbound shipments are bundled in the originating depot by source area, and are rebundled at hubs by delivering spoke (depot). Inbound and outbound services are typically operated by the same local fleets. However, hub-and-spoke networks operated with sub-contracted depots still have room for improvement by addressing the following issues:

- The organizational heterogeneity of depots (different policies, handling of shipment information, etc.) impairs interoperability. This also covers reliability, the timing of sharing information, etc.

- Multiple-participant networks can suffer from information not being available in a given place at the right time (ie data flow may lag behind the material flow), resulting in sub-optimal decisions and inefficient or unbalanced use of transportation assets. This is critical if the same vehicles perform both depot-to-hub and hub-to-depot traffic and leave/arrive by the same schedule.

- In a large network, the amount of data presents its own challenges in addition to the challenges associated with surveying longer time ranges to detect trends or patterns. Massive but low-level information must be prepared for decisions, forecasts, etc. The implicit knowledge of demand dynamics and dependencies is currently largely unexploited.

- Confidentiality issues may also arise if members collaborating within the network compete on other services. While participants strive to disclose a safe minimum of information, it may be possible to employ privacy-preserving techniques.

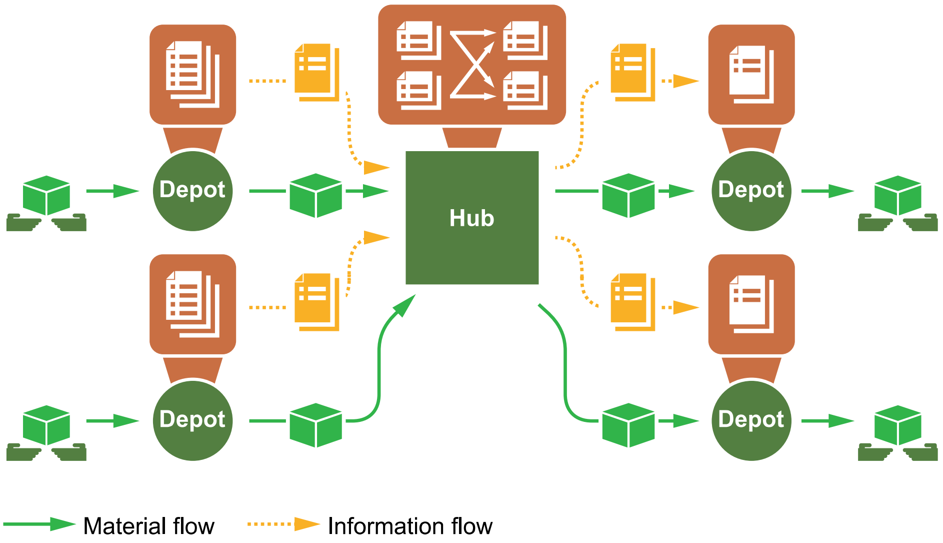


Figure 1: In complex logistics networks, data accompanying the shipments are just as important as the material stream itself—in certain cases, shipment data have to be filtered, reformulated, or need to precede the material flow for adequate allocation of resources. This is not a trivial challenge and has to comply with constraints as privacy of network participants.

- Making participation attractive may conflict with overall network efficiency. This hindrance is well-known in other related domains, such as supply chain management (SCM), and has been addressed by various risk-sharing and profit-sharing contracting policies.

The European R&D project ADVANCE addresses several of the aforementioned challenges with a solution framework providing data transformation, filtering, model building and prediction tools for operational-level decision support. Working closely with the UK-based company, Palletways, gives the project consortium an excellent background for in-depth field studies, and allows objective testing of pilot applications. The project will deliver:

- Runtime environment and flow editor. Dispatching and processing massive low-level information requires an efficient runtime environment. To this end, the ADVANCE project employs the reactive approach whose asynchronous, event-oriented characteristics guarantee the lean usage of computational resources. The framework uses data models specifically targeting the logistics sector. Nevertheless, these are kept flexible enough for deployment in other networks. IT staff adopting the framework to specific use cases will have a flow editor at hand. This is, in essence, a graphical programming interface for laying out connections between functional blocks before run-time use, with advanced productivity and verification features (e.g. type probing of data channels and type inference).

- Exploiting low-level information. Today's logistics networks are characterized by massive but distributed low-level data where relevant information may be implicit or remain localized. This apparent challenge bears much potential which the ADVANCE project aims to harness. Machine learning techniques are employed to build process models and extract relevant information, making it available for operational decisions or on-demand forecasts.

- Decision support on the operational level. ADVANCE envisages decision support functionalities in various forms, appropriate for different situations. Where needed a model-based prediction produces forecasts whilst in other cases, warnings are issued to call the operator's attention to the need for intervention. The evolvability of decision support is highly dependent on the way in which the results fit into the operator's mental context; in other words, it is best if the way the suggestion is presented "makes sense". In the ADVANCE project, cognitive models of human reactions are employed to adapt to the nature of human comprehension, attention and misconceptions.

Started in 2010, the three-year R&D project ADVANCE is financed within the EU 7th Framework Programme, and has already successfully completed its first year. Led by SZTAKI of Hungary, the consortium also includes academic members from the UK (Aston University) and the Netherlands (RUG), an Italian IT development company (TTS), and UK-based logistics company Palletways, the latter hosting pilot applications and field tests. Generic parts of the project's output will be freely available to the public.

Link:
http://www.advance-logistics.eu/

Please contact:

Elisabeth Ilie-Zudor
SZTAKI, Hungary
Tel: +36 1 279 6195
E-mail: ilie@sztaki.hu

Digital Material Appearance: The Curse of Tera-Bytes

by Michal Haindl, Jiří Filip and Radomír Vávra

Real surface material visual appearance is a highly complex physical phenomenon which intricately depends on incident and reflected spherical angles, time, light spectrum and other physical variables. The best current measurable representation of a material appearance requires tens of thousands of images using a sophisticated high precision automatic measuring device. This results in a huge amount of data which can easily reach tens of tera-bytes for a single measured material. Nevertheless, these data have insufficient spatial extent for any real virtual reality applications and have to be further enlarged using advanced modelling techniques. In order to apply such expensive and massive measurements to a car interior design, for instance, we would need at least 20 such demanding material measurements.

Within the Pattern Recognition department of UTIA, we have built a high precision robotic gonioreflectometer [2] (see Figure1). The setup consists of independently controlled arms with camera and light. Its parameters such as angular precision (to 0.03 degrees), spatial resolution (1000 DPI), and selective spatial measurement qualify this gonioreflectometer as a state-of-the-art device. The typical resolution of an area of interest is around 2000 x 2000 pixels, each of which is represented by at least 16-bit floating point values to achieve reasonable representation of high-dynamic-range visual information. The memory requirements for storage of a single material sample amount to 360 giga-bytes per colour channel. Three RGB colour channels require about one tera-byte. More precise spectral measurements with a moderate visible spectrum (400-700nm) sampling further increase the amount of data to five tera-bytes or more. Such measurements (illustrated in Figure 2) allow us to capture the very fine meso-structure of individual entities comprising the material.

Such enormous volumes of visual data inevitably require state-of-the-art solutions for storage, compression, modelling, visualization, and quality verification. Storage technology is still the weak link, lagging behind recent developments in data sensing technologies. Our solution is a compromised combination of fast but overpriced disk array and slow but cheap tape storage. The compression and modelling steps are integrated due to our visual data representation based on several novel multidimensional probabilistic models. Some of these models (using either a set of underlying Markov random fields or probabilistic mixtures) are described in ERCIM News 81 [1]. These models [2] allow unlimited seamless material texture enlargement, texture restoration, huge unbeatable appearance data compression (up to 1:1000 000) and even editing or creation of novel material appearance data. They require neither storing of original measurements nor any pixel-wise parametric representation.

A further problem is that of physically correct material visualization, because there are no professional systems which allow rendering of such complex data. Therefore, we were forced to develop the novel Blender plugin for the purpose of realistic material appearance model mapping and rendering.

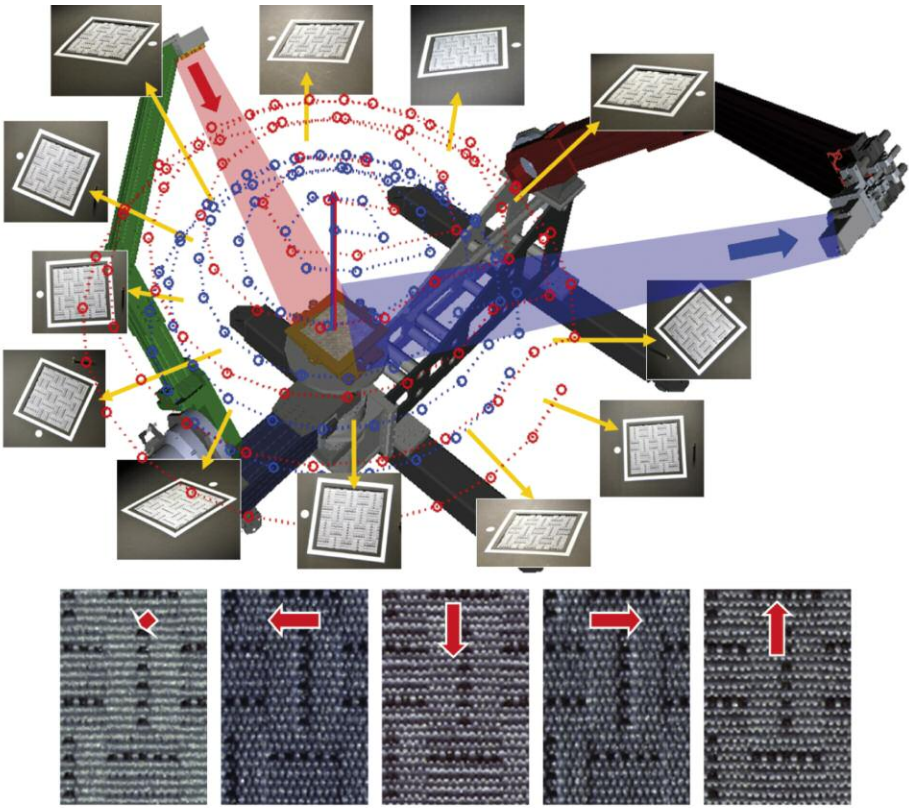


Figure 1: Appearance measurement principle. The total number of images is equal to every possible combination of red dots (camera directions) and blue dots (light directions). Below is an example of how the visual appearance of fabric is dependent upon illumination direction. Fabric is illuminated from above, left, top, right and below, respectively.

Blender is a free open source 3D graphics application for creation 3D models, visualizations and animations and is available for all major operating systems under the GNU General Public License. Visual quality verification is another difficult unsolved problem which we tackle using applied psychophysically validated techniques.

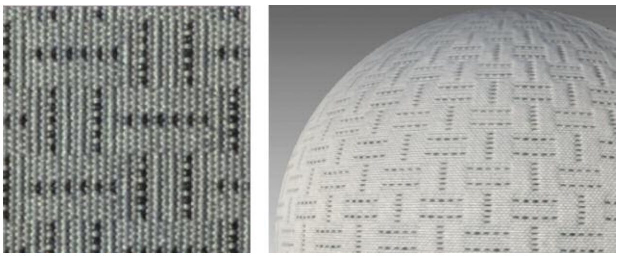


Figure 2: Detail of measured fabric material (left) and its large-scale mapping on a sphere (right).

These precise visual measurements are crucial for a better understanding of human visual perception of real-world materials. This is the key challenge not only for efficient and simultaneously convincing representations of visual information but also for further progress in visual scene analysis. Due to their tricky and expensive measurements only a small amount of data of low quality are available so far. Therefore, we provide benchmark material measurements for image and computer graphics research purposes on our web server listed below.

Link:
http://btf.utia.cas.cz/

References:

[1] Haindl M., Filip J., Hatka M.: Realistic Material Appearance Modelling . ERCIM News (2010), 81, pp.13-14
[2] Haindl, M., Filip, J. Advanced textural representation of materials appearance, Proc. SIGGRAPH Asia '11 (Courses), pp. 1:1-1:84, ACM.

Please contact:

Michal Haindl
CRCIM (UTIA), Czech Republic
Tel: +420 266052350
E-mail: haindl@utia.cas.cz

(No)SQL Platform for Scalable Semantic Processing of Fast Growing Document Repositories

by Dominik Ślęzak, Krzysztof Stencel and Son Nguyen

Large volumes of scientific data require specific storage and indexing solutions to maximize the effectiveness of searches. Semantic indexing algorithms use domain knowledge to group and rank sets of objects of interest, such as publications, scientists, institutes, and scientific concepts. Their implementation is often based on massively parallel solutions employing NoSQL platforms, although different types of semantic processing operations should be scaled with respect to the growing volumes of scientific information using different database methodologies [1].

SONCA (Search based on ONtologies and Compound Analytics) platform is developed at the Faculty of Mathematics, Informatics and Mechanics of the University of Warsaw. It is part of the project 'Interdisciplinary System for Interactive Scientific and Scientific-Technical Information' (www.synat.pl). SONCA is an application based on a hybrid database framework, wherein scientific articles are stored and processed in various forms. SONCA is expected to provide interfaces for intelligent algorithms identifying relations between various types of objects [2]. It extends typical functionality of scientific search engines by more accurate identification of relevant documents and more advanced synthesis of information. To achieve this, concurrent processing of documents needs to be coupled with ability to produce collections of new objects using queries specific for analytic database technologies.

SONCA's architecture comprises four layers (Figure 1). User Interface receives requests in a domain-specific language, rewrites them into the chains of (No)SQL statements executed iteratively against the Semantic Indices and – in some cases – the contents of SoncaDB, and prepares answers in a form of relational, graph or XML structures that can be passed to external reporting and visualization modules.

The Semantic Indices are periodically recomputed by Analytic Algorithms basing on continuously growing contents of SoncaDB. SoncaDB stores articles (and other pieces of information) acquired from external sources in two formats: XML structures extracted for each single document using structural OCR, and subsets of tuples corresponding to documents' properties, parts, single words and relationships to other objects, populated across data tables in a relational database.

The roles of Analytic Algorithms and the relational subpart of SoncaDB are two examples of SONCA's innovation. Additional storage of full information about articles in a tabular form gives developers of Analytic Algorithms a choice between relational, structural and mixed methods of data access, data processing, and storage of the obtained outcomes. However, for millions of documents, we should expect billions of rows. Hence, the underlying RDBMS technologies need to be very carefully adjusted.

The increase of volumes of tuples with each document loaded into SONCA is faster than one might expect. Each article yields entities corresponding to its authors, bibliography items, and areas of science related to thematic classifications or tags. These entities are recorded in generic tables as instances of objects (such as scientist, publication, area and so on) with some properties (e.g. scientist's affiliation or article's publisher) and links to a document from which they were parsed (including information about the location within a document that a given article was cited, a given concept was described and so on). Instances are grouped into classes corresponding to actual objects of interest (for instance: bibliographic items parsed from several documents may turn out to be the same article) using (No)SQL procedures executed over SoncaDB. Analytic Algorithms are then adding their own tuples corresponding, for instance, to heuristically identified semantic relations between objects that improve the quality of search processes.

Owing to the rapid growth in the volume of data we use technologies based on intelligent software rather than massive hardware. We chose Infobright's analytic RDBMS engine (www.infobright.com) to handle the relational subpart of SoncaDB because of its performance on machine-generated data originating from sources such as sensor measurements, computer logs or RFID readings. Although SONCA gathers contents created by humans, the way in which they are parsed and extended makes them more similar to machine-generated data sets. We also use carefully selected solutions for other database components such as the structural subpart of SoncaDB (MongoDB is employed here because of its flexibility of enriching objects with new information) and the Semantic Indices (outputs of Analytic Algorithms can be stored in Cassandra, Lucene or PostgreSQL, for example, depending on how they are used by User Interface modules).

The performance and quality tests undertaken so far on over 200K full-content articles resulting in 300M tuples confirm SONCA's scalability [3], which should be investigated not only by means of data volume but also ease of adding new types of objects that may be of interest for specific groups of users. The relational data model employed within SoncaDB enables smooth extension of the set of supported types of objects with no need to create new tables or attributes. It is also prepared to deal on the same basis with objects acquired at different stages of parsing (e.g. concepts derived from domain ontologies vs. concepts detected as keywords in loaded texts) and with different degrees of information completeness (e.g. fully available articles vs. articles identified as bibliography items elsewhere). However, as already mentioned, the crucial aspect is freedom of choice between different data forms and processing strategies while optimizing Analytic Algorithms, reducing execution time of specific tasks from (hundreds of) hours to (tens of) minutes.

References:

[1] R. Agrawal et al.: The Claremont Report on Database Research. Commun. ACM 52(6), 56-65 (2009).
[2] R. Bembenik et al. (eds.): Intelligent Tools for Building a Scientific Information Platform. Springer (2012).
[3] D. Ślęzak et al.: Semantic Analytics of PubMed Content. In: Proc. of USAB 2011, 63-74.

Please contact:

Dominik Ślęzak
Institute of Mathematics, University of Warsaw, Poland
and Infobright Inc., Canada
E-mail: slezak@mimuw.edu.pl

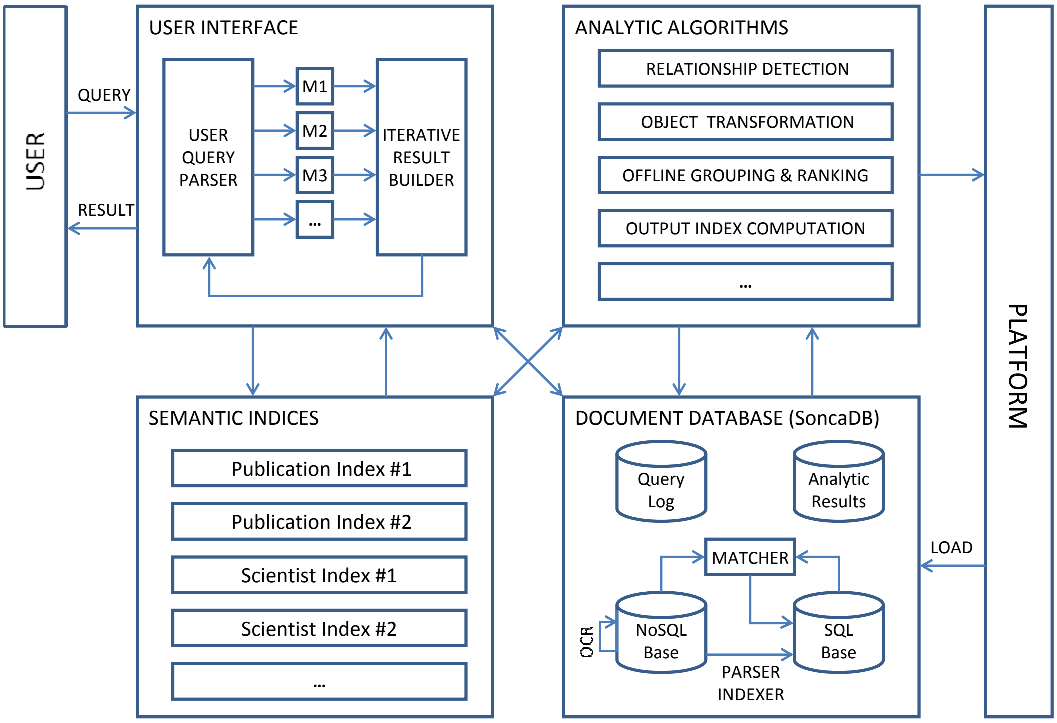


Figure 1: Major layers of SONCA's architecture.

The “SHOWN” Platform - Safeguarding Architectural Heritage Using Wireless Networks

by Paolo Barsocchi and Maria Girardi

Regular surveys of the structure of World Heritage buildings are essential to ensure their conservation. The aging of building materials, long-term ground subsidence, and environmental vibrations are all possible causes of the deterioration of old constructions. The potentially disastrous effects of severe seismic events are an additional factor teaching us the importance of prevention. When the constraints of architectural conservation are very strict, prevention mainly means monitoring.

While great efforts have been made to monitor the surfaces of ancient monuments, the level of the technology applied to control their “structural health” is still generally quite low. A typical programme of structural monitoring simply involves technicians using classical optical instruments periodically carrying out a series of measurements. More recently, procedures have been developed to test the structural health of ancient buildings by analysing their dynamic response to natural or artificial vibrations. Such procedures are recognized as a good way to test the state of conservation of a building, and are also important aids in identifying when interventions are necessary. They consist in taking regular measurements via wired acceleration and displacement sensors, which are removed after usage. Lately, microwave interferometry technology has been used to measure the vibration of buildings at great distances. Unfortunately, these techniques generally involve high maintenance costs and thus make the continuous acquisition of the large amounts of data necessary for effective monitoring impossible. In addition, wired sensors are too invasive and aesthetically unacceptable for widespread application to the architectural heritage.

In this context, Wireless Sensor Network (WSN) technology can make an important contribution by providing an economical and relatively non-invasive instrument for real-time structural monitoring of the well-being of buildings and monuments. We envisage their employment in a not too distant future in the monitoring of entire historic areas on a large-scale, facilitating the management of maintenance operations and prompt interventions in the case of an emergency. The installation of a WSN during the construction or the restoration of a building and its connection to a central database could become an ordinary or even compulsory operation. For these reasons, the structural monitoring of ancient buildings using wireless sensor network technology seems very promising. It could provide continuous observation and real-time feedback, allowing engineers to reconfigure the network, as necessary. WSN-based technology also offers the added benefits of low visual impact and low maintenance costs.

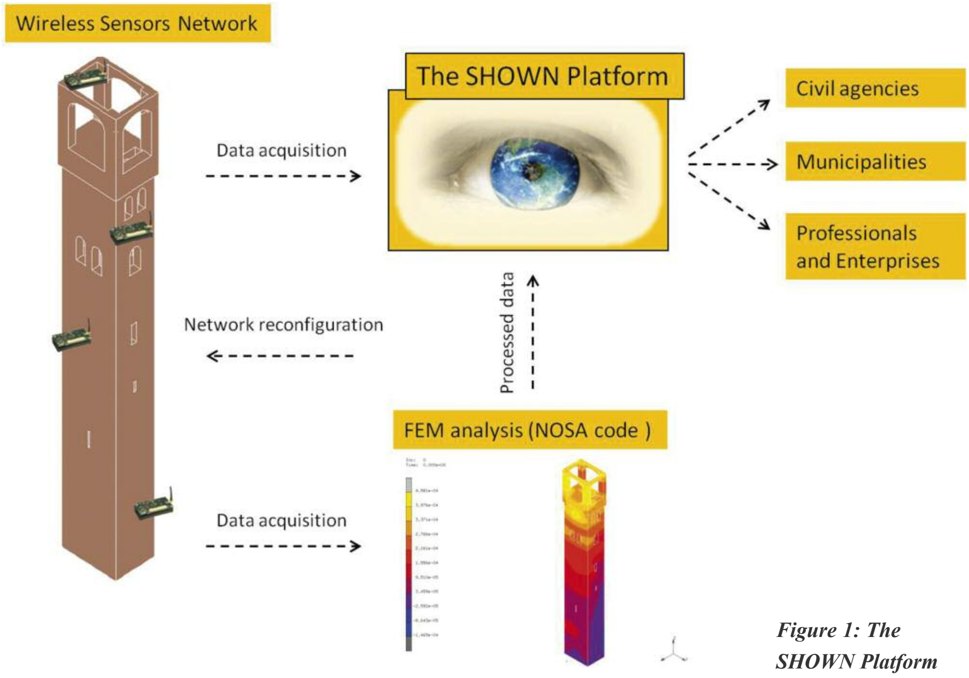


Figure 1: The SHOWN Platform

Up to now, in the field of cultural and architectural heritage, WSNs have been mainly used to monitor large archaeological areas or some interior parameters of ancient buildings and museums, such as moisture, temperature and fire. Applications of WSN technology to the structural monitoring of ancient buildings so far have been limited to few example applications and research projects. At the moment, several issues remain to be investigated, such as the number and location of the sensors used to acquire data and the maximisation of WSN lifetime. While energy consumption is a well-known problem in WSN applications, the optimisation of the number and location of sensors is a new challenge in these technologies, which typically involve a large number of redundant sensors.

The SHOWN platform, depicted in Figure 1, will move in this direction. The research is conducted by the Mechanics of Materials and Structures Laboratory and the Wireless Sensor Networks Laboratory of ISTI–CNR and a recent breakthrough is the development of wireless networks able to communicate through a minimum number of sensors, taking into account both structural and architectonic constraints on the one hand and radio signal propagation constraints on the other. Considerable attention is also being given to energy optimisation of the network, with particular regard to the high–frequency sampling sensors, such as accelerometers.

All structural parameters are set and verified using the NOSA-ITACA code, a finite element software developed at ISTI-CNR to simulate the static and dynamic behaviour of ancient masonry.

Links:
[1] The Mechanics of Materials and Structures Laboratory: http://www.isti.cnr.it/research/unit.php?unit=MMS

[2] The Wireless Networks Laboratory (WN Lab) : http://www.isti.cnr.it/research/unit.php?unit=WN

[3] http://www.nosaitaca.it

Please contact:
Paolo Barsocchi and Maria Girardi, ISTI-CNR, Italy
E-mail: paolo.barsocchi@isti.cnr.it, maria.girardi@isti.cnr.it

Cite-as-you-Write

by Kris Jack, Maurizio Sambati, Fabrizio Silvestri, Salvatore Trani, Rossano Venturini

When starting a new research activity, it is essential to study related work. Traditional search engines and dedicated social networks are generally used to search for relevant literature. Current technologies rely on keyword based searches which, however, do not provide the support of a wider context. Cite-as-you-write aims to simplify and shorten this exploratory task: given a verbose description of the problem to be investigated, the system automatically recommends related papers/citations.

Recommender systems for scientific papers have received much attention during the last decade and some novel approaches have been experimented. We propose an innovative contextual search engine tool to address this problem. The tool exploits several aspects of the scientific literature ecosystem including an intriguing social angle derived from data provided by the popular Mendely platform. The objective is to provide pointers to existing literature given a description of the study the researcher is undertaking.

We began by building a baseline method to retrieve literature related to a fragment of text describing the research concept. This consisted of a system taking as input a textual description of the research idea and returning the most similar (according to a similarity measure) papers in the literature. To evaluate the similarity between the query and documents the “cosine similarity” metric was used. This metric calculates how many terms are in common between the query and each document and weights the terms (in query and documents) by an importance score.

We subsequently refined the baseline strategy by adopting a “Learning to Rank” approach. Within this approach the similarity between queries and documents is computed via a metric that is “learned” from samples input to a machine learning algorithm. The main difference between a search engine and Cite-as-you-write consists in how the queries are formulated: search engines are usually optimized to answer keyword-based queries, our system extracts a context from a long description of the research problem provided by the scientist.

The system consists of three modules:
- Crawler, which builds and maintains the repository of scientific literature.
- Indexer, which processes the scientific papers and builds an index on their most representative terms.
- Query processor, a specialized search engine with an advanced ranking algorithm designed for the task of retrieving related work from a detailed context.

The data used to build and evaluate our system consists of about 500 thousand computer science papers including their citations. The data was kindly provided to us by Mendeley.

The index represented under the form of an inverted index contains only the most representative terms for each paper. This trade-off in coverage, keeps down the size of the index. Fortunately. our experiments show that the loss due to reduced coverage is limited, as scientific publications usually focus on a few specific topics represented by a small number of important terms.

Following the typical approach of learning-to-rank-based retrieval systems, the ranking phase consists of two steps:
1. A cosine similarity ranking based on the title and the abstracts of the papers (ie, the baseline method)
2. A ranking function that combines all the features we have collected from the data

The second step works by adopting a technique known as similarity learning, which consists in exploiting sample data to find correlations between a set of features and a target variable. The learning method adopted is a Random Forest ensemble which uses our features based on text similarity, paper publication date, the citation network (i.e. PageRank and absolute number of citations) and the Mendeley environment (ie popularity of papers in user libraries with some normalizations: the importance of the user, defined in the social network as the number of links, and by the number of items in the library, reducing the weight when a user has lots of papers in their library). Random Forest is a very simple but powerful machine learning tool which builds several decision trees from various random samplings of the data. The result is the average of the results returned by each tree. The same strategy as that adopted in democratic voting systems.

Experiments shown in Figure 1 show the improvements on the test data over the baseline system with variations in the number of trees (x-axis).

We used the normalized discounted accumulative gain (nDCG) evaluation metric to measure the effectiveness of a ranking, ie how near relevant documents were to the top of the result lists [1]. The baseline considered (the black line) is the cosine similarity based ranking. The improvement provided by our tool on the test set with respect to the base line is summarized in Figure 2.

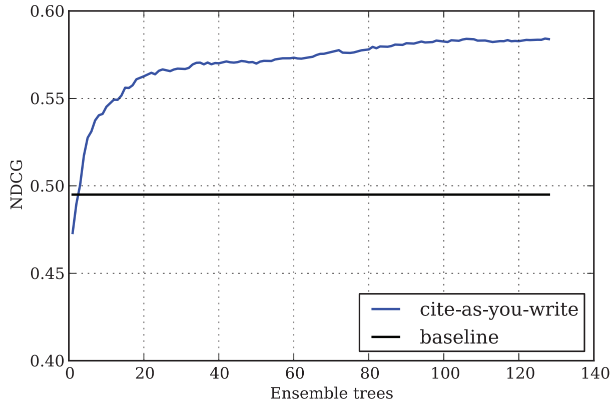


Figure 1: improvements on the test data over the baseline system with variations in the number of trees (x-axis)

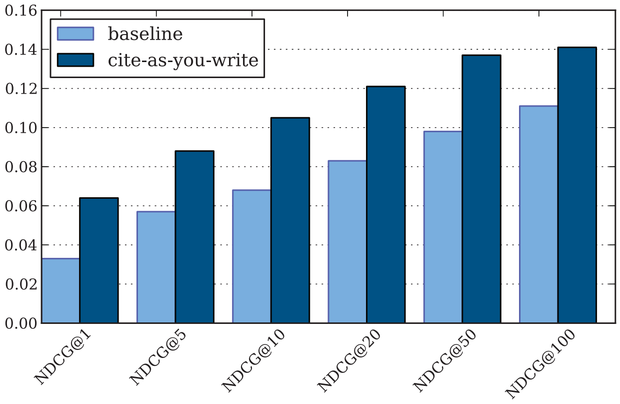


Figure 2: test ranking results

Learning to rank and recommender systems are emerging fields in computer science and have been applied to real world problems in a number of ways. We have used and adapted cutting edge technologies to build a powerful tool for a more natural scientific paper recommendation with respect to traditional search engines. We believe context-driven search can be a better alternative to keyword based search for next generation web.

Links:
Cite-as-you-write (prototype):
http://vinello.isti.cnr.it/cite-as-you-write/
Tool used for learning to rank:
https://bitbucket.org/duilio/mltool
Mendeley: http://www.mendeley.com/
Random Forests:
http://stat-http://www.berkeley.edu/users/breiman/RandomForests/cc_home.htm

Reference:

[1] Kalervo Järvelin and Jaana Kekäläinen. 2002. Cumulated gain-based evaluation of IR techniques. ACM Trans. Inf. Syst. 20, 4 (October 2002), 422-446.
DOI=10.1145/582415.582418
http://doi.acm.org/10.1145/582415.582418

Please contact:

Maurizio Sambati and Salvatore Trani
ISTI-CNR, Italy
E-mail. maurizio.sambati@gmail.com,
trani.salvatore@gmail.com

ElasticSSI: Self-optimizing Metacomputing through Process Migration and Elastic Scaling

by Philip Healy, John Morrison, Ray Walshe

Single System Image (SSI) systems allow a collection of discrete machines to be presented to the user in the guise of a single virtual machine. Similarly, Infrastructure-as-a-Service (IaaS) interfaces allow one or more physical machines to be presented to the user in the guise of a collection of discrete virtual machines. Creating an SSI instance from a pool of virtual resources provisioned from an IaaS provider affords the ability to leverage the "on-demand" nature of IaaS to quickly and easily adjust the size of the resource pool. With ElasticSSI we propose to automate this adjustment process through the application of elastic scaling. The automation of the scaling process will result in systems that are self-optimizing with respect to resource utilization; virtual resources are allocated and released based the value of system load metrics, which in turn are dependent on the resources allocated. The scaling process is transparent to the end user as the SSI system maintains the illusion of interacting with a single Linux OS instance.

Single System Image (SSI) is a metacomputing paradigm where a number of discrete computing resources are aggregated and presented to the user via an interface that maintains the illusion of interaction with a single system. Although the interface chosen could represent any one of a number of resource types at varying levels of abstraction, the term SSI has become synonymous with the abstraction of clusters of machines at the kernel level. The greatest advantage of the process migration approach to parallelism is that in most cases the user does not need to modify application code in order to parallelize a particular workload. However, performance gains are only possible if the workload in question is amenable to process migration. Suitable applications include long-running shell scripts, place and route, and 3D rendering.

The rise of virtualization in recent years has lead to the popularization of Infrastructure-as-a-Service (IaaS). Under the IaaS model, resources (such as virtual machine instances and block devices) can be provisioned as needed by the end user. Although the IaaS model can be applied to physical hardware, it is typically used to provide access to virtualized resources. The IaaS provisioning process becomes more interesting when it is automated, i.e., the system modifies itself by allocating and deallocating resources in response to rises and falls in demand, a technique referred to as elastic scaling. Virtualization and SSI solve similar problems in that both divorce the resources available to an OS from the physical hardware.

To date the SSI approach has been applied mostly to physical compute clusters and grids where it has gained some popu-

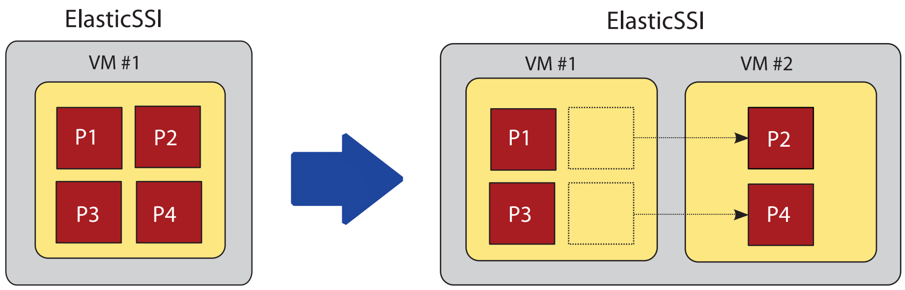


Figure 1: As the load on an ElasticSSI instance increases, new virtual machines are added and processes are migrated.

larity but has never emerged as a mainstream technique. The reasons underlying the lack of uptake for SSI were examined in a series of blog posts by Greg Pfister, a former IBM Distinguished Engineer and cluster computing specialist (see link below). Pfister identified availability during OS updates, application compatibility issues, cognitive inertia, and matching the parallelism of a workload to the underlying computing resources available. We are cognizant of these issues and postulate another: barrier to entry. The creation and administration of compute clusters is an activity outside the core competency of many organizations, and so tends to be performed only in situations where the need is acute. Taking a chance on exotic, experimental systems such as SSI with associated kernel patching appears to be a bridge too far for most organizations.

The ElasticSSI project, led by Philip Healy and John Morrison of University College Cork in conjunction with Ray Walshe of Dublin City University, intends to address these issues by creating an SSI implementation that builds on the IaaS model to provide elastic scaling. We believe that the barrier to entry for SSI can be removed by making our implementation available as a Platform-as-a-Service (PaaS) offering; users will be able to create an instance of our system at the push of a button and experiment from there for evaluation purposes. The PaaS approach also solves the issue of OS updates as users simply create a new ElasticSSI instance whenever an OS upgrade is required. We do not propose to take any significant steps to address application compatibility or workload matching as these go against the simplicity that characterizes the SSI approach; in our view, if the user's anticipated workload is not a good fit from the outset then attempting to force a fit will rapidly lead to diminishing returns. The issue of cognitive inertia is, we believe, closely related to the barrier to entry issue; solving the latter would go a long way towards addressing the former.

Related work includes MIT’s Factored Operating System (fos), which is an attempt to create a scalable cloud-based SSI operating system from scratch. Although this approach has considerable merit, the scope of the fos project is much more ambitious than the simple “process migration with elastic scaling” approach advocated here. In contrast, ElasticSSI will be a minimal implementation based on an existing Linux distribution. This is in line with our goal of reaching as wide an audience as possible and encouraging experimentation. In light of this, we have decided that the best starting point is one of the existing Linux-based SSI implementations, of which there are several. These are typically implemented as a set of kernel patches that implement the required OS modifications along with user-mode tools and libraries that implement complementary non-kernel functionality. Existing implementations include MOSIX, OpenMOSIX, LinuxPMI, Kerrighed and OpenSSI. A final decision on which implementation to adopt will be taken once a thorough evaluation of each has been completed.

SSI implementations do not require homogeneity across the machines that make up the underlying resource pool. Similarly, IaaS providers typically offer a range of virtual machine types, with some types being tailored for compute-intensive workloads (by increasing the number and type of available cores) and others similarly geared towards memory-intensive workloads (by increasing the amount of memory available). An active area of research will be to introduce heterogeneity into the resource pool by allocating virtual machine instances of varying types based on the value of metrics such as system-wide CPU load and memory utilization.

References:

- Rajkumar Buyya, Toni Cortes and Hai Jin, Single System Image, International Journal of High Performance Computing Applications, 15 (2): 124.
- David Wentzlaff, Charles Gruenwald III, Nathan Beckmann, Kevin Modzelewski, Adam Belay, Lamia Youseff, Jason Miller, and Anant Agarwal, A Unified Operating System for Clouds and Manycore: fos, 1st Workshop on Computer Architecture and Operating System co-design (CAOS), Jan 2010.

Link:
http://perilsofparallel.blogspot.com/2009/01/multi-multi-core-single-system-image.html

Please contact:

Philip Healy
Irish Centre for Cloud Computing and Commerce, Ireland
Tel: +353 21 4205935
E-mail: p.healy@cs.ucc.ie

Event Report

Workshop on "Global Scientific Data Infrastructures: The Findability Challenge"

by Costantino Thanos

It is well-known that the scientific world is creating an unimaginably vast amount of rapidly increasing digital data. Among the members of the academic research community, there is growing consensus that e-science practices should be congruent with open-science and that scientific data should be freely accessible. However, in a networked open-science world, a big challenge faced by researchers is findability. Findability means the ease with which data/information/knowledge and tools/services for specific purposes can be discovered, and takes into account relevant aspects of the attributes, context and provenance of the data, the functionality and deployability of the tools and services, and profiles and goals of the searcher, etc. On the contrary, the current Internet search paradigm is characterized by a lack of context, with search being conducted independently of data provenance, professional profiles, and work goals. Enabling findability is thus of paramount importance for the next generation of global research data infrastructures.

A Workshop, held in May 2012 in Taormina, Italy, and organized by ISTI-CNR, aimed at investigating the findability challenge. The Workshop was organized around 14 invited talks offered by internationally recognized scientists working in the areas of databases, information retrieval, knowledge representation, and data infrastructures.

It was first suggested that the findability challenge could be considered as understanding how we can bring together information about a single subject (such as a topic or an entity) scattered across different sources (e.g., web sites, social sites, news feeds) in order to gain more complete and possibly more accurate knowledge about this process.

A number of topics were then examined in-depth. It was felt that one of the main emerging needs with big data applications is data exploration, i.e. examining big data sets searching for interesting patterns without a precise idea of what is being looked for. In this respect, adaptive query processing, ie, adaptive indexing, adaptive data loading, etc., was considered to be an appropriate technology that can help data management systems and scientists to avoid expensive actions until they are absolutely certain that these actions are going to pay off. The notion of mega-modeling, i.e. modeling a certain aspect/system, and then creating new services/models by combining existing ones in a principled way was also considered important. Some of the talks addressed the relevance of concepts such as semantics, modeling, and ontologies for findability. It was recognized that the design of efficient semantic query answering services remained a real challenge. Another findability challenge that was given considerable attention was scalability. When it comes to scalability, there are two sides of the coin: (a) scalability wrt data: ie, keeping up with the amount of online data, and (b) scalability wrt knowledge: the richer, more complete, more accurate the knowledge we seek the more difficult it is to acquire it. There was general agreement that it was important to be able to discover not just data but also data tools and services, ie, to enable the automated location of data tools and services that adequately fulfill a given research need.

Finally, the inadequacy of the current database technology to address the requirements of science was emphasized and some research directions for a more effective scientific data management in the context of a data intensive research were illustrated.

The scientific program of the Workshop was coordinated by Costantino Thanos (ISTI-CNR) and Yannis Ioannidis (Univ. of Athens). Further details and online versions of many of the invited talks can be found on the workshop website

More information:
http://datachallenges.isti.cnr.it/

Call for Participation

SICS Software Week in September 2012

Welcome to SICS Software Week – three of SICS' most popular open conferences gathered together in one week in September! Three days of seminars, exhibitions and mingle – a perfect reload after summer vacations.

- Sept. 10: Internet of Things Day
- Sept. 11: Cloud Day
- Sept. 12: Multicore Day

SICS organizes a number of events on selected technology hot topics every year. These three days, now put together in one powerful software week, have in the past each attracted around 300 people from Swedish and international industry, academia, and the public sector. It is a nice opportunity to share ideas and discuss the latest trends with researchers and industry in an informal setting.

SICS is located in Kista, Stockholm, Sweden. The talks are in English.

More information:
details about the speakers, program and registration at http://www.sics.se/ssw2012

HCI International 2013

Las Vegas, USA, 21-26 July 2013

The 15th International Conference on Human-Computer Interaction, HCI International 2013, will be held jointly with the affiliated Conferences:

- Human-Computer Interaction thematic area
- Human Interface and the Management of Information thematic area
- 10th International Conference on Engineering Psychology and Cognitive Ergonomics
- 7th International Conference on Universal Access in Human-Computer Interaction
- 5th International Conference on Virtual, Augmented and Mixed Reality
- 5th International Conference on Cross-Cultural Design
- 5th International Conference on Online Communities and Social Computing
- 7th International Conference on Augmented Cognition
- 4th International Conference on Digital Human Modeling
- 2nd International Conference on Design, User Experience and Usability
- 1st International Conference on Distributed, Ambient and Pervasive Interactions - NEW
- 1st International Conference on Human Aspects of Information Security, Privacy and Trust - NEW

Please visit the Conference website for further information on each thematic area / conference, including topics and Program Board members.

HCI International 2013 is expected to attract over 2,000 participants from all over the world. The program will feature, among others, pre-conference half-day and full-day tutorials, parallel sessions, poster presentations, an opening session with a keynote address, and an exhibition.

The Conference Proceedings will be published by Springer in a multi-volume set. Papers will appear in volumes of the LNCS and LNAI series. Extended Poster abstracts will be published in the CCIS series. All volumes will be available on-line through the SpringerLink Digital Library, readily accessible by all subscribing libraries around the world, and will be indexed by a number of services including EI and ISI CPCI-S.

The best paper of each of the Affiliated Conferences / Thematic Areas will receive an award. Among these best papers, one will be selected to receive the golden award as the Best HCI International 2013 Conference paper. Finally, the Best Poster extended abstract will also receive an award.

Links:
HCI International 2013,
http://www.hcii2013.org/
HCI International Conference Series,
http://www.hci-international.org/

Please contact:
Constantine Stephanidis
General Chair, HCI International Conf.
Email: cs@ics.forth.gr

3DStereoMedia 2012

Liège, Belgium, 3-6 December 2012

3D StereoMedia 2012 is the the fourth edition of the European 3D Stereo Summit for Science, Technology and Digital Art. The multi-faceted event includes the 3D Academy workshop, the IC3D scientific conference co-sponsored by IEEE, a 3D film market, a 3D film festival, a professional conference, and an exhibition. On 6 December 2012 in the evening, the "Lumière Statuettes" will be awarded in 14 categiries during the I3DS-Europe Awards gala.

3DSM is one of the few places in the world where 3D experts can meet, exchange ideas, conceive projects, and conclude deals.

More information:
http://www.3dstereomedia.eu

ICT Proposers' Day

Warsaw, 26-27 September 2012

The ICT Proposers' Day is a unique networking opportunity to build partnerships and projects targetting the new Information and Communication Technologies Work Program me of the EU's 7th Framework Programme for 2013.

Everyone who is interested in responding to calls for proposals for R&D projects in the field of Information and Communication Technologies. It is an exceptional occasion to meet potential partners from academia, research institutes, business and industry, SMEs and government actors from all over Europe. More than 2000 participants are expected to attend.

The event will provide:

- first-hand information from European Commission officials on the ICT Work Programme 2013, offering around 1.5 billion euro of EU funding
- answers to questions related to the upcoming calls for proposals
- an opportunity to present and discuss your project idea during one of the networking sessions
- a platform for exchanging ideas and finding right partners to form project consortia
- guidance on how to present a successful proposal.

The event is organised by by the Directorate-General for Communications Networks, Content and Technology of the EC, in cooperation with the Polish Ministry of Science and Higher Education and the Polish National Contact Point.

More information:
http://ec.europa.eu/information_society/events/ictproposersday/2012/index_en.htm

W3DevCampus: Next Mobile Web Courses Dates Published

W3C is pleased to announce the schedule of upcoming mobile Web courses, from September 2012 until April 2013. The three mobile Web online training courses are:

- “Mobile Web 1: Best Practices”
- “Buenas Prácticas en Web Móvil” (which is the Spanish version of "Mobile Web 1")
- “Mobile Web 2: Programming Web Applications”

During these courses, participants learn the basic coding techniques and design principles for the mobile Web, understand the specifics of developing Web applications for the mobile environment, learn the latest HTML5 and Javascript APIs that are actually usable in real-world environments, and much more.

The next mobile Web courses will start as soon as on 3 September 2012, so make sure to register soon. All the course details and other informations are available on the W3DevCampus Web site.

Since their inception in 2008, the mobile Web courses have attracted over 1500 students and with each new module, the W3DevCampus program attracts more and more attention. Developed by the MobiWebApp EU project team, the courses builds on the output of the various expert groups operating within W3C.

About W3DevCampus

W3C offers high-quality online training for Web developers, worldwide. All courses follow a similar format: weekly modules consisting of lectures and links to further resources, followed by practical exercices such as quizzes and/or assignments. A discussion forum allows participants to discuss the course with each other and with the instructors.

More information:
http://www.w3devcampus.com/

Forthcoming ERCIM Working Group Workshops

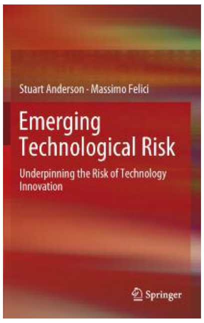
- 28-29 August 2012, Paris: 16th International Workshop on Formal Methods for Industrial Critical Systems (FMICS 2012) held in conjucntion with FM’12
- 6 September 2012, Brussels: Research Challenges in Software Complexity, in conjunction with ECCS’12
- 8 September 2012, Edinburgh: Joint Workshop on Compositional Modelling and Analysis of Quantitative Systems. The workshop will take place immediately following the CONCUR Conference, Newcastle-upon-Tyne 4-7 September 2012
- 10-12 September 2012, Pisa, Italy: The ERCIM Working Group Security and Trust Management will hold its 2012 workshop in conjunction with the European Symposium on Research in Computer Security (ESORICS) 2012
- 25 September 2012, Magedburg, Germany: The ERCIM Working Group Dependable Embedded Systems will organise a workshop at SAFECOMP 2012
- December 1-3, 2012, Oviedo, Spain: 5th International Conference of the ERCIM WG on Computing & Statistics.

http://www.ercim.eu/activity/f-events

Books

Stuart Anderson, Massimo Felici

Emerging Technological Risk: Underpinning the Risk of Technology Innovation


Stuart Anderson · Massimo Felici
Emerging Technological Risk
Underpinning the Risk of Technology Innovation
Springer

Classes of socio-technical hazards allow a characterization of the risk in technology innovation and clarify the mechanisms underpinning emergent technological risk. Emerging Technological Risk provides an interdisciplinary account of risk in socio-technical systems including hazards.

Addressing an audience from a range of academic and professional backgrounds, Emerging Technological Risk is a key source for those who wish to benefit from a detail and methodical exposure to multiple perspectives on technological risk. By providing a synthesis of recent work on risk that captures the complex mechanisms that characterize the emergence of risk in technology innovation, Emerging Technological Risk bridges contributions from many disciplines in order to sustain a fruitful debate.

Emerging Technological Risk is one of a series of books developed by the Dependability Interdisciplinary Research Collaboration funded by the UK Engineering and Physical Sciences Research Council.
Springer 2012, http://www.springer.com/978-1-4471-2142-8

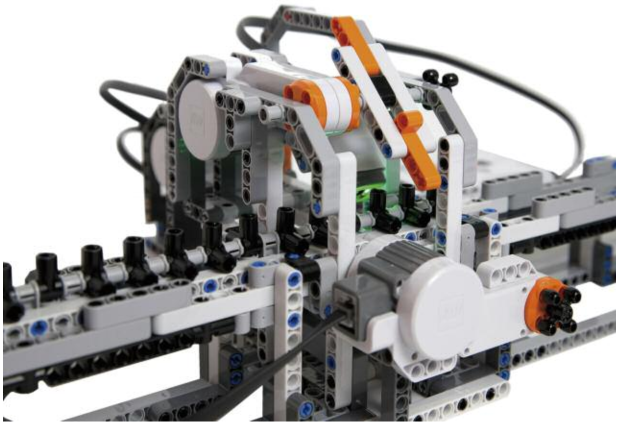
Computer Networks Journal Special Issues

The Computer Networks journal has recently published two special issues:
Since Peer-to-Peer systems are the largest contributor to Internet traffic, it is important to measure this traffic and so to be able to optimize its flow. This was the reason for collecting papers for a special issue entitled “Measurement-based Optimization of P2P Networking and Applications”, Volume 56, Number 3, published in February 2012.

As our networks grow they become more complex. A second special issue of Computer Networks, also in Volume 56, Number 3, is entitled “Complex Dynamic Networks: Tools and Methods” and is dedicated to capturing the state-of-the-art and recent advances on tools and methods for characterizing, analyzing, and visualizing these networks.
http://www.journals.elsevier.com/computer-networks/

CWI researchers built LEGO Turing machine in Turing Year

CWI researchers Jeroen van den Bos en Davy Landman from CWI built a working Turing machine out of LEGO bricks. This machine shows the working of a computer, based on the ideas of the famous British mathematician Alan Turing in 1936. Turing (1912-1954) is known as the founding father of computer science and artificial intelligence, and as Enigma code breaker in World War II. He was one of the few people of his time that realized the influence that computers could have on society. Van den Bos and Landman used a single LEGO Mindstorms NXT set to build the Turing machine, to make it easy to reproduce for those interested. It is on display until 6 October at the temporary exhibition ‘Turing’s Legacy’ at CWI, organized during the Turing Year 2012 to commemorate Alan Turing’s centenary on 23 June. The exhibition at CWI covers Alan Turing’s work and shows how his ideas are applied in contemporary research.



Reading the memory. Source: CWI.

More information:
http://www.legoturingmachine.org/
video: http://vimeo.com/44202270

GoalRef: FIFA Approves Intelligent Goal from Fraunhofer

Goal or no goal? In response to this question, world football association FIFA wants to use technical assistance in the future. In its meeting of Thursday 5th July 2012 the International Football Association Board (IFAB), the body which determines the laws of the game, approved both goal-line technologies GoalRef and Hawk-Eye.This approval is subject to a final installation test at each stadium before the systems can be used in “real” football matches, in accordance with the FIFA Quality Programme for GLT.

FIFA, football’s world governing body, and a member of the IFAB has decided to use both goal-line technology systems from GoalRef and Hawk-Eye at the FIFA Club World Cup in Japan in December this year. The GoalRef system was developed by researchers from the Fraunhofer Institute for Integrated Circuits IIS. “The technology works in a similar way to that of the theft pr otection of a department store,” explained René Dünkler, spokesman of the GoalRef project. Ten antennae behind the goalpost and crossbar create and monitor a weak magnetic field. As soon as the ball nears the goal-line the field is influenced by thin spools in the football. A processor is then able to determine, by means of the antenna signal, whether the ball fully crossed the goal-line or not. “GoalRef is a bit like an invisible curtain which hangs behind the crossbar and the goal-line. As soon as the ball fully passes through this ‘curtain’, it is recognised as a goal,” says Ingmar Bretz, project head of GoalRef. The system then automatically sends this information in real time via encoded radio signals to the referees whose special wrist watches display the result visually and by means of vibration.

GoalRef reliably determines whether or not the whole ball has passed the goal line. Photo: Fraunhofer IIS/Kurt Fuchs.

More information:
http://www.iis.fraunhofer.de/en/bf/ln/referenzprojekte/goalref/index.jsp

ICT Foresight for Southeast European Countries

The "FORSEE - Regional ICT Foresight exercise for Southeast European countries" project targets ICT RTD policy reform in the South-eastern Europe (SEE) region, proposing a focused effort on introducing a foresight culture in the region, which is necessary in order to accelerate socio-economic growth in participating countries, striving to meet the challenges of the global networked economy and to participate on equal footing in the European Research Area.

The FORSEE initiative aims to introduce a sustainable mechanism for ICT Foresight in the region, attempting to tackle the absence of a regular process applied for technological future orientation and research policy review.

The project lasts until end of 2013 and is funded by the South East Europe Transational Cooperation Programme. The Industrial Systems Institute of RC Athena joined the project in early 2012. Currently, the consortium is in the phase of identifying the ICT themes and topics of interest for a pilot regional foresight exercise.

More information: http://www.forsee.eu/

ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development, in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

ERCIM is the European Host of the World Wide Web Consortium.

Austrian Association for Research in IT
c/o Österreichische Computer Gesellschaft
Wollzeile 1-3, A-1010 Wien, Austria
http://www.aarit.at/

Consiglio Nazionale delle Ricerche, ISTI-CNR
Area della Ricerca CNR di Pisa,
Via G. Moruzzi 1, 56124 Pisa, Italy
http://www.isti.cnr.it/

Czech Research Consortium
for Informatics and Mathematics
FI MU, Botanicka 68a, CZ-602 00 Brno, Czech Republic
http://www.utia.cas.cz/CRCIM/home.html

Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
http://www.cwi.nl/

Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
http://www.fnr.lu/

FWO
Egmontstraat 5
B-1000 Brussels, Belgium
http://www.fwo.be/

FNRS
rue d'Egmont 5
B-1000 Brussels, Belgium
http://www.fnrs.be/

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
http://www.ics.forth.gr/

Fraunhofer ICT Group
Anna-Louisa-Karsch-Str. 2
10178 Berlin, Germany
http://www.iuk.fraunhofer.de/

Institut National de Recherche en Informatique
et en Automatique
B.P. 105, F-78153 Le Chesnay, France
http://www.inria.fr/

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and
Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

I.S.I. - Industrial Systems Institute
Patras Science Park building
Platani, PATRAS, Greece, 265 04
http://www.isi.gr

Portuguese ERCIM Grouping
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal

Polish Research Consortium for Informatics and Mathematics
Wydział Matematyki, Informatyki i Mechaniki,
Uniwersytetu Warszawskiego, ul. Banacha 2, 02-097 Warszawa, Poland
http://www.plercim.pl/

Science and Technology Facilities Council,
Rutherford Appleton Laboratory
Harwell Science and Innovation Campus
Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
http://www.scitech.ac.uk/

Spanish Research Consortium for Informatics and Mathematics,
D3301, Facultad de Informática, Universidad Politécnica de Madrid,
Campus de Montegancedo s/n,
28660 Boadilla del Monte, Madrid, Spain,
http://www.sparcim.es/

Swedish Institute of Computer Science
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

Swiss Association for Research in Information Technology
c/o Professor Abraham Bernstein, Ph.D., Department of
Informatics, University of Zurich, Binzmühlestrasse 14,
CH-8050 Zürich
http://www.sarit.ch

Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
http://www.sztaki.hu/

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
http://www.cs.ucy.ac.cy/

Technical Research Centre of Finland
PO Box 1000
FIN-02044 VTT, Finland
http://www.vtt.fi/

Order Form

If you wish to subscribe to ERCIM News ***free of charge*** *or if you know of a colleague who would like to receive regular copies of ERCIM News, please fill in this form and we will add you/them to the mailing list.*

Send, fax or email this form to:
ERCIM NEWS
2004 route des Lucioles
BP 93
F-06902 Sophia Antipolis Cedex
Fax: +33 4 9238 5011
E-mail: contact@ercim.eu

Data from this form will be held on a computer database.
By giving your email address, you allow ERCIM to send you email

I wish to subscribe to the

☐ ***printed edition*** ☐ ***online edition (email required)***

Name:

Organisation/Company:

Address:

Postal Code:

City:

Country

E-mail:

You can also subscribe to ERCIM News and order back copies by filling out the form at the ERCIM Web site at http://ercim-news.ercim.eu/