

# Federico Maggi

Politecnico di Milano

Dipartimento di Elettronica e Informazione

Room T.16, Building 18

Via Ponzio 34/5

I-20133 Milano- Italy

T +39 02 2399 4009

F +39 02 2399 3411

email: [fmaggi@elet.polimi.it](mailto:fmaggi@elet.polimi.it)

www: <http://home.dei.polimi.it/fmaggi/>

Born: June 13, 1982 in Vimercate, Milano - Italy

Nationality: Italian

---

*Last updated on February 26, 2010<sup>1</sup>*

## Areas of specialization

Computer security (anomaly-based intrusion detection, alert correlation, web application firewalls, anti-forensics).

## Current position

1/2010- *Contract Researcher*, Dip. di Elettronica e Informazione  
Politecnico di Milano, Italy

1/2007- *Ph.D. Candidate*, Dip. di Elettronica e Informazione  
Politecnico di Milano, Italy

## Association Memberships

2008- IEEE - *Institute of Electrical and Electronics Engineers*  
Student member.


## Education

2007-2009 PH.D. in Computer Engineering. *Politecnico di Milano*. Italy.  
THESIS: *Integrated Detection of Anomalous Behavior of Computer Infrastructures*  
ADVISOR: Prof. S. Zanero.

9/2008-6/2009 *Visiting Scholar*, Computer Security Laboratory, Dept. Computer Science  
University of Santa Barbara, California, United States

2004-2006 M.S. in Computer Engineering. *Politecnico di Milano*. Italy.

---

<sup>1</sup>  Si autorizza al trattamento dei dati per finalità di reclutamento/selezione ex Decreto Legislativo 196/2003.

- THESIS: *Efficacia ed integrazione di sistemi di anomaly detection*  
 ADVISOR: Prof. G. Serazzi. GRADE: *Highest honors.*
- 2001-2004 M.S. in Computer Engineering. *Politecnico di Milano.* Italy.  
 THESIS: *Portable DBMS: analysis, specification and prototype implementation of a JDBC layer*  
 ADVISOR: Prof. C. Bolchini. GRADE: *101/110.*
- 1995-2001 DIPLOMA di Maturità Scientifica. *Liceo Scientifico-Tecnologico.* Vimercate, Italy.  
 GRADE: *86/100.*

#### SCHOOL PARTICIPATION

- 8/2007 SWING *Second PhD and Researchers International School on Security for Wireless Networking.* Bertinoro, Italy.
- 3/2007 EWSCS *13th Estonian Winter School in Computer Science.* Palmse, Estonia.
- 8/2008 SWING *Third PhD and Researchers International School on Security for Wireless Networking.* Bertinoro, Italy.

## Teaching Activities

#### TEACHING ASSISTANTSHIPS

- 2005-2007 060012: Informatica 1 (*lab. tutor*). Prof. C. Bolchini.  
 B.S. in Computer Engineering. *Politecnico di Milano.* Italy
- 2007-2008 060012: Informatica 1 (*lab. supervisor*). Prof. C. Bolchini.  
 B.S. in Computer Engineering. *Politecnico di Milano.* Italy
- 2007-2008 070728: Sicurezza degli impianti informatici (*teaching assistant*). Prof. G. Serazzi.  
 M.S. in Computer Engineering. *Politecnico di Milano.* Italy
- 2007- 072614: Impianti informatici (*teaching assistant*). Prof. G. Serazzi.  
 B.S./M.S. in Computer Engineering. *Politecnico di Milano.* Italy
- 2008-2009 061187: Impianti di Elaborazione (*teaching assistant*). Prof. S. Zanero.  
 B.S. in Management Engineering. *Politecnico di Milano.* Italy
- 2009-2010 060016: Informatica B (*lab. tutor*). Prof. A. Campi.  
 B.S. in Mechanic Engineering. *Politecnico di Milano.* Italy
- 2009-2010 085630: Sistemi Informativi (*teaching assistant*). Prof. E. Capra.  
 B.S. in Management Engineering. *Politecnico di Milano.* Italy

#### CO-ADVISOR IN LAUREA THESIS

- 2008-2009 *E. Gressi*, Apprendimento e simulazione dell'attività di un utente mediante l'utilizzo di modelli semi-markoviani nascosti.  
 Politecnico di Milano, Italy.
- 2007-2008 *M. Michellini*, Kernel auditing su Linux 2.6 in formato OpenBSM.  
 Politecnico di Milano, Italy.
- 2006-2007 *P. Testa*, Valutazione automatica delle performance di sistemi di anomaly detection.  
 Politecnico di Milano, Italy.

- 2006-2007 *C. Magni*, Analisi e test automatizzati di sistemi di anomaly detection network-based. Politecnico di Milano, Italy.
- 2006-2007 *M. Debiasi, M. Falsitta*, Reingengerizzazione ed ottimizzazione di un sistema di anomaly detection host based. Politecnico di Milano, Italy.

## Professional Activities

- 2007- IT TEACHER. *Verbano Informatica* of Paolo Garlassi. Verbania, Italy. Professional teaching activity.
- 2001-2005 FREELANCE IT CONSULTANT. Network design, deployment, and testing; custom system sizing; customer support. MAIN TECHNOLOGIES: xDSL, TCP/IP, Linux, Windows NT Server, Windows.
- 2001- FREELANCE WEB DEVELOPER (*Secure-Web.it*). Concept, design, implementation, deployment and maintenance of small to medium websites. MAIN TECHNOLOGIES: Web standards, AJAX, CSS, XHTML, XML, modern frameworks.
- 2005- IT JOURNALIST *Infomedia*. Ponsacco, Pisa, Italy. Short publications about bleeding-edge technologies for the Web.
- 2005- JUNIOR IT CONSULTANT. *SecureNetwork s.r.l.*. Milano, Italy. Web application penetration testing and vulnerability assessment.
- 9-12/2000 SUMMER JOB. *ST Microelectronics*. Agrate, Milano, Italy. Production operations.

## Main Awards

- 2007 M.S. THESIS. One of the 6 best theses. *Premio tesi CLUSIT*. Milano, Italy. <http://clusit.it/archivio.htm#premiotesi2007>
- 2007 M.S. THESIS. Candidate at the final selection. *Premio laurea Accenture*. Milano, Italy.

## Research Activity

My research interest is *computer security*. In particular, my activity focuses on the use of anomaly-based *intrusion detection* techniques to protect operating systems' kernel and web applications. Also, I study the problem of automatic and reliable accuracy and performance testing of intrusion detection systems.

### BEHAVIORAL MODELS FOR ANOMALY DETECTION

The continuous evolution of attacks and malware against computer infrastructures is unveiling the limits of the classical approaches to intrusion detection. The *anomaly based* approach to the detection of anomalous activities currently needs for effective, accurate, and robust models for characterizing normal behaviors to correctly detect deviations (e.g., attacks, intrusions, malware, etc.). We proposed an unsupervised host-based IDS to detect anomalies into the operating system calls of the Linux kernel. We defined anomaly detection models for each of the system calls and used

machine learning techniques to capture canonical behavior of processes and flag anomalies according to proper thresholds. The whole system needs no prior knowledge input and has a good signal to noise ratio.

Also, we proposed a so called “hybrid host-based IDS”. To detect anomalies on system call sequences and arguments it combines our machine learning detection techniques with another approach based on finite state machines and data flow models. Such a result shows how the accuracy of an hybrid IDS can be enhanced by using system call arguments and data flow models.

We also proposed an original idea to detect attempts of circumventing forensic analysis using anti-forensic techniques. In particular, we shown how the use of host anomaly detection tools can reveal unwanted malicious activity of active processes. Our detection techniques are capable of detecting malicious behavior even if the attacker is stealth and/or uses anti-forensic tools.

The main results of the research have been published in (? ? ), and (? ).

### Detection of Attacks Against Web Applications

Research on (machine) learning based intrusion detectors has always suffered from the lack of realistic, unbiased data for both training and accuracy evaluation.

This problem, shared with many other research fields, is magnified by the complexity of todays’ threat scenario that is populated by powerful, collaborative and well-organized attackers (e.g., botnets). It not only demands for complex techniques to model the legit activity of protocols, making the intrusion detection task inherently more difficult; more importantly, the systems to protect have very loose boundaries (e.g., web applications, services, web platforms), thus, generating or gathering good training data is a very difficult task.

We recently developed an approach for detecting attacks against the web browser, the web application and the database (? ). In addition, we developed an adaptive approach that allows to perform training of a learning-based detector in an automated fashion, especially in the case of web applications subject to frequent changes (? ) and in the case of freshly deployed web applications, when scarce training data is available (? ).

### ALERT CORRELATION

The majority of currently available anomaly detectors only works on isolated systems or local applications. Instead, we strongly believe that the integration of both information sources (i.e., observed activities) and revealed anomalies (i.e., alerts), is the first step toward the development of enhanced anomaly detection techniques. Apparently trivial, the integration and the correlation of anomaly and misuse detectors is one of the open problems in the intrusion detection research.

We propose an unsupervised approach based on a set of statistical tests, and we prove that our criteria work well on a simplified correlation task, without requiring complex configuration parameters. We also focused on the aggregation of IDS alerts, an important component of the alert fusion process. We proposed to exploit fuzzy measures and fuzzy sets to design simple and robust alert aggregation algorithms. We published our main results in (? ), and (? ).

### OTHER RESEARCH INTERESTS

During my PhD I did some minor research in the field of theoretical computer science. In particular, I investigated the membership problem of rational trace languages.

### Trace Languages

Mazurkiewicz’s Traces provide a simple and powerful framework to model *concurrency* by extending string languages to capture the *dependencies* among instructions. The *theory of traces* has been applied to the real-world problem of *automatic instruction parallelization* with promising results. Traces can

also be exploited for security-related problems, such as anomaly detection, where abstract models are learned and leveraged to detect deviations. A notable example is the use of system calls sequences to formalize all the possible behavioral profiles of programs. In this case, traces allow an alternative, compact and lightweight model representation with benefit of both space and effectiveness. While a string language is used to describe the syntax of computer programs, a trace language capture the *dependencies* among instructions. From this point of view, trace languages extend the expressive power of string languages allowing the specification of simple *semantic* constraints. However, in most of the real-world applications, *time* is the main concern. Indeed, the demand for fast recognition algorithms makes traces *unsuitable* for real applications so far.

We focus on the *Membership Problem* (MP), which plays a key role in real-world applications where classical and expensive dynamic programming techniques are used. The result of our research have been published in a technical report (?).

## Publications

### JOURNAL PUBLICATIONS

- [1] F. Maggi, M. Matteucci, and S. Zanero, “Detecting intrusions through system call sequence and argument analysis,” *IEEE Transactions on Dependable and Secure Computing*, vol. 99, no. PrePrints, 2008, [PDF \(preprint\)](#).
- [2] F. Maggi, S. Zanero, and V. Iozzo, “Seeing the invisible: forensic uses of anomaly detection and machine learning,” *SIGOPS Oper. Syst. Rev.*, vol. 42, no. 3, pp. 51–58, 2008, [PDF](#).
- [3] F. Maggi, M. Matteucci, and S. Zanero, “Reducing false positives in anomaly detectors through fuzzy alert aggregation,” *Information Fusion*, 2009, [PDF](#).

### CONFERENCE AND WORKSHOP PROCEEDINGS

- [4] Claudio Criscione, Federico Maggi, Guido Salvaneschi, and Stefano Zanero. Integrated detection of attacks against browsers, web applications and databases. In *European Conference on Computer Network Defence*, Milano, Italy, 2009. [PDF](#).
- [5] Alessandro Frossi, Federico Maggi, Gian Luigi Rizzo, and Stefano Zanero. Selecting and improving system call models for anomaly detection. In Ulrich Flegel and Michael Meier, editors, *Detecting Intrusions through System Call Sequence and Argument Analysis*, Lecture Notes in Computer Science, Como, Italy, 2009. Springer. [PDF](#).
- [6] Federico Maggi, William Robertson, Christopher Kruegel, and Giovanni Vigna. Protecting a moving target: Addressing web application concept drift. In Engin Kirda and Davide Balzarotti, editors, *Recent Advances in Intrusion Detection*, St Malo, Brittany, France, 2009. Springer. [PDF](#).
- [7] Federico Maggi and Stefano Zanero. On the use of different statistical tests for alert correlation - short paper. In Christopher Krügel, Richard Lippmann, and Andrew Clark, editors, *Recent Advances in Intrusion Detection*, volume 4637 of *Lecture Notes in Computer Science*, pages 167–177, Surfer’s Paradise, Queensland, Australia, 2007. Springer. [PDF](#).
- [8] William Robertson, Federico Maggi, Christopher Kruegel, and Giovanni Vigna. Effective anomaly detection with scarce training data. In *Annual Network & Distributed System Security Symposium*, San Diego, CA, USA, March 2010.

## TECHNICAL REPORTS

In addition to the technical report cited above I also published a survey ( ? ) on probabilistic record-matching techniques, although this does not overlap with any of my main research activities.

- [9] Federico Maggi. Specification and evaluation of an efficient recognizer for rational trace languages. Technical Report TR-2008-23, Politecnico di Milano, 2008. [PDF](#).
- [10] Federico Maggi. A survey of probabilistic record matching models, techniques and tools. Technical Report TR-2008-22, Politecnico di Milano, 2008. [PDF](#).

## TALKS

- Feb 25, 2010 Detecting Anomalous Behaviors in Computer Infrastructures ([PDF](#)).  
*Fondazione Bruno Kessler, Trento - Italy.*
- Jan 21, 2010 “Just-in-Time” Training of Anomaly Detectors ([PDF](#)).  
*Vrije Universiteit Amsterdam - The Netherlands.*
- Jul 21, 2009 Training with (almost) no data - Addressing Training Issues In Modern Intrusion Detection ([PDF](#)).  
*Politecnico di Milano - Italy.*

## IT & Programming Skills

I tend to adapt and extend my knowledge as demanded by the project I am working on. However, I have some skills and preferences in terms of technologies I use every day.

### Programming languages

1. Python
2. PHP
3. Java
4. C

### Scripting languages

1. Bash

### Web frameworks

1. Django
2. JavaScript

### 3. PrototypeJS

4. jQuery
5. CakePHP

### Markup languages

1. XHTML
2. CSS
3. XML
4. XSD

### Typesetting

1. L<sup>A</sup>T<sub>E</sub>X
2. XeTeX

### Databases systems

1. PostgreSQL
2. SQLite
3. MySQL

### Operating systems

1. Mac OS X
2. Ubuntu
3. Debian/GNU Linux
4. FreeBSD
5. Windows

## Languages

*Italian* (native speaker).

*English* (fluent).

## References

PROFESSOR GIUSEPPE SERAZZI - [serazzi@elet.polimi.it](mailto:serazzi@elet.polimi.it)

104 Edificio 15  
Dipartimento di Elettronica e Informazione  
Politecnico di Milano  
Via Ponzio 34/5  
I-20133 Milano, Italy

PROFESSOR STEFANO ZANERO - [zanero@elet.polimi.it](mailto:zanero@elet.polimi.it)

114 Edificio 15  
Dipartimento di Elettronica e Informazione  
Politecnico di Milano  
Via Ponzio 34/5  
I-20133 Milano, Italy

PROFESSOR CHRISTOPHER KRÜGEL - [chris@cs.ucsb.edu](mailto:chris@cs.ucsb.edu)

1117 Engineering I  
Department of Computer Science  
University of California, Santa Barbara  
Santa Barbara, CA 93106, United States

PROFESSOR GIOVANNI VIGNA - [vigna@cs.ucsb.edu](mailto:vigna@cs.ucsb.edu)

2159 Engineering I  
Department of Computer Science  
University of California, Santa Barbara  
Santa Barbara, CA 93106, United States