Sam L. Thomas *Doctoral Researcher in Information Security*

WWW	https://exo.st	GitHub	https://github.com/xorpse
Google Scholar	https://goo.gl/eNBqdY	LinkedIn	https://linkedin.com/in/xorpse

Profile

A highly (self-)motivated doctoral researcher in software and systems security at the University of Birmingham with expertise in reverse engineering and security analysis of binary code; consistently awarded best student in his degree programme in consecutive years throughout his bachelor's degree.

At the age of 14 he was already experienced in his first programming language: `C` which quickly motivated an interest in the low-level aspects of computer architecture – learning `x86` assembly language and going on to reverse engineer proprietary software protections for fun. This later progressed into more academic interests in compiler and programming language design, influencing him to develop a complete compiler front-end and code synthesiser for the `C` programming language as the topic of his bachelor's thesis.

During the course of his bachelor's degree he cultivated an interest in the more formal aspects of computer science such as program analysis, proof and verification. He also gained significant experience in a number of programming languages including `OCaml`, `C++`, `Java` and `Python`. In the final year of his bachelor's degree, he was offered a Ph.D. studentship to work with Flavio Garcia and Tom Chothia, specifically developing automated program analysis techniques for detecting backdoors within embedded device firmware.

His Ph.D. has seen him independently develop a number of tools for semi-automated backdoor detection: incorporating aspects of machine learning and traditional program analysis and furthering his assembly language repertoire to include `ARM` and `MIPS`. Through this work, he has authored two published papers at well respected academic conferences.

Outside of academia, his expertise resulted in him being employed as an independent consultant for Huawei Technologies, which he undertook alongside his studies. He has also taken a leading role in the University of Birmingham's Capture The Flag team: AFiniteNumberOfMonkeys (2014 – 2015, the term where they ranked highest), where he further developed his reverse engineering skills and acquired skills in both web and binary exploitation.

Experience

EMSEC @ IRISA — Rennes, France
Visiting Researcher — *Mar '18 – Mar '18*
This role was undertaken as a Short Term Scientific Mission (STSM) at IRISA working with Clémentine Maurice under a grant from COST for the IC1403 action – CRYPTACUS (Cryptanalysis of ubiquitous computing systems). The work focused on developing a testbed for performing microarchitectural research.

Huawei Technologies — Remote
Independent Consultant — *Jan '17 – Oct '17*
This role was undertaken as joint work with Flavio Garcia and Tom Chothia. The project consisted of producing a number of reports and holding live workshops; the topics broadly covered analysis of Android, router and IoT firmware backdoors as well as methods of auditing and detection.

Software & Systems Security Group (S3) @ EURECOM — Biot, France
Visiting Researcher — *Jun '17 – Jun '17*
This role was undertaken as a Short Term Scientific Mission (STSM) at EURECOM working with Aurélien Francillon under a grant from COST for the IC1403 action – CRYPTACUS (Cryptanalysis of ubiquitous computing systems). The work focused on developing a framework for understanding backdoors in the context of their detection and deniability. The process involved learning a number of state of the art tools and frameworks for automated program analysis and vulnerability detection such as `angr` and `Driller`.

University of Birmingham — Birmingham, United Kingdom
Teaching Assistant (Secure Programming) — *2016 – 2017, 2017 – present*
This position involved holding weekly lab sessions: providing students with hands-on experience in discovering program vulnerabilities, writing working exploits and patching services. The course is a masters-level course on secure programming and covers all aspects of building secure programs. It incorporates topics such as secure password storage, mitigating SQL injection, CSRF and XSS attacks as well as buffer overflows, type confusion and race conditions.

Teaching Assistant (Introduction to Computer Security) *2015 – 2016*

This position involved marking student assignments and providing relevant feedback. The course is an undergraduate-level course covering the basics of computer security including: cryptography, exploit development, protocols and reverse engineering.

Teaching Assistant (Foundations of Computer Science) *2012 – 2013*

This position was undertaken during my bachelor's degree working alongside Ph.D. student demonstrators, the role involved holding a weekly office hour and demonstrations. The course is an undergraduate-level course which covers the basics of computer science from a theoretical standpoint: algorithms, complexity theory and reasoning about program correctness. It also covered more advanced topics such as hand-written proofs by structural induction on basic functional programs written in OCaml.

Recommendations/references available on request.

Education

University of Birmingham — Birmingham, United Kingdom

Doctor of Philosophy (Ph.D.) in Information Security *2014 – 2018 (expected)*

Thesis title: *BaDSeED: Backdoor Detection Systems for Embedded Devices*

During the course of the Ph.D. programme, numerous diverse topics have been explored, ranging from traditional program analysis methods such as data flow analysis and more modern topics such as symbolic execution to machine learning. Much of the work produced has seen these aforementioned topics combined in novel ways.

For example, the tool `HumIDIFy` – the result of the first strand of research undertaken – is based upon first using semi-supervised learning techniques to identify common functionality classes of binary software, which then drives targeted static analysis passes which, in turn identifies unexpected functionality. These static analysis passes are encoded in a domain specific language – `Binary Functionality Description Language` (also developed as part of this project), which allows for specifying program properties in a high-level, readable format.

The second strand of research saw the development of another tool, `Stringer` which identifies key static data and associated branching constructs by means of a ranking metric. The static data ranked the highest is that when successfully matched against acts as a trigger to executing otherwise unreachable parts of the program. These otherwise unreachable program parts are synonymous with backdoor payloads. As part of this work, algorithms to automatically identify static data comparison functions were developed; which perform with high precision and minimal false positives.

Aside from technical contributions, a more formal discussion of backdoor implementation and design has been explored which addresses the limitations in automated backdoor (and vulnerability) detection as well as assigning accountability by quantifying the deniability of particular backdoor implementation techniques.

Bachelor's Degree in Computer Science (First Class Honours) *2011 – 2014*

Thesis title: *MetaC: Enriching the C Programming Language with DSL Extensions*

The thesis topic involved creating a compiler front-end (lexer, parser and AST representation) for `C` (specifically `C99`). In addition to handling standard `C`, the front-end additionally supported parsing arbitrary embedded domain specific languages which were implemented as compiler plugins. The compiler itself synthesised the domain specific languages to `C`. In addition to the compiler front-end, the project also contained an `OCaml`-based API for defining new domain specific languages. The project was composed of in excess of `20k` lines of OCaml.

Select modules undertaken during the degree programme include: Operating Systems with `C` and `C++`, Principles of Programming Languages, Team Programming, Compilers and, Computer Systems & Architecture.

King Edward VI College — Stourbridge, United Kingdom

A Levels in: *2009 – 2011*

- Physics (A*)
- Mathematics (A)
- Computing (A)
- Film Studies (A)

Full academic transcript available on request.

Awards & Scholarships

University of Birmingham — Birmingham, United Kingdom

Best Final Year Student in Degree Programme *2014*

Computer Science (School) Prize *2014*

Best 2nd Year Student in Degree Programme *2013*

Science Undergraduate Prize *2013*

Best 1st Year Student in Degree Programme *2012*

School of Computer Science Excellence Scholarship *2011*

BrumHack (Hackathon) — Birmingham, United Kingdom

Prize Winner *2015*

Created a proof-of-concept solution: *deCAPTCHA* using image recognition API from Clarifai and NLP techniques (porter stemming) to successfully break Google's reCAPTCHA challenges.

King Edward VI College — Stourbridge, United Kingdom

Baylies Educational Foundation Award *2011*

Governor's Leaving Scholarship Award *2011*

Trustee's Prize (for Mathematics) *2010*

Publications

Stringer: Measuring the Importance of Static Data Comparisons to Detect Backdoors and Undocumented Functionality

22nd European Symposium on Research in Computer Security (ESORICS'17)

Sam L. Thomas, Tom Chothia and Flavio D. Garcia

This paper describes Stringer, a tool which implements techniques developed to detect potential backdoors and undocumented functionality in embedded device firmware. It employs a novel static analysis method which identifies key static data and branching constructs. It has been used to identify a number of backdoors within various types of consumer products such as routers and DVR devices.

HumIDIFy: A Tool for Hidden Functionality Detection in Firmware

14th International Conference on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA'17)

Sam L. Thomas, Flavio D. Garcia and Tom Chothia

This paper describes HumIDIFy, a tool for large-scale analysis of embedded device firmware. It uses a hybrid of machine learning and static analysis to identify anomalous functionality within binaries extracted from embedded device firmware. It is effective in identifying backdoors with very low runtime requirements.

Abstracts and full papers available on request.

Talks

2018

The Internet of Backdoors

Invited talk to present research on backdoors and backdoor detection at the SoSySec seminar (Rennes, France).

2017

Stringer: Measuring the Importance of Static Data Comparisons to Detect Backdoors and Undocumented Functionality

Conference talk to present paper content at the 22nd European Symposium on Research in Computer Security (ESORICS'17; Oslo, Norway).

HumIDIFy: A Tool for Hidden Functionality Detection in Firmware

Conference talk to present paper content at the 14th International Conference on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA'17; Bonn, Germany).

HumIDIFy: A Tool for Hidden Functionality Detection in Firmware

Invited talk to present extended conference talk and discussion of paper content and tool implementation (HumIDIFy) with the S3 group (EURECOM; Biot, France).

2016

Firmware Backdoors

Invited talk at RSSB (Rail Safety and Standards Board) to present (from a high-level perspective) the threats posed by backdoors to embedded systems and also participated in panel of experts on related topics on the next generation of rail networks.

Malware Detection: Firmware Backdoors

Invited lecture for Anonymity, Privacy AND Cybercrime course at University of Birmingham on malware detection with a focus on detecting backdoors by using machine learning to guide traditional program analysis methods.

Slides available on request.

Courses & Summer Schools

2016

Computer Aided Analysis of Cryptographic Protocols Bucharest, Romania

The topic was the foundations of computer aided verification for cryptographic systems. The talks covered topics such as theorem proving and type inference as well as a number of tools: Proverif, Tamarin, F*, Easycrypt, which were applied to both basic security primitives and much larger protocols such as TLS.

2015

International School on Foundations of Security Analysis and Design Bertinoro, Italy

The talks covered various topics related to formal security analysis of protocols and programs such as: security of multi-party computations, privacy in the Internet of the future, cryptographic and probabilistic programming, automated security proofs by contradiction and systems security (specifically botnet take-overs and SROP).

Summer School on real-world crypto and privacy Šibenik, Croatia

The school involved basic and advanced talks on various topics in cryptography, from theory to implementation and attacks: side channel and fault injection attacks and countermeasures, provable security, physical attacks, lightweight and real-world cryptography, secure implementations in hardware and software, and the mathematics of public-key cryptography.

2014

Cryptography I Online

This online course offered by Coursera covered topics such as secure key-exchange, symmetric key cryptography and public key cryptography, attacks on those constructs and various security properties. It also offered hands-on programming exercises dealing with implementation of protocols, systems and attacks. Course was completed with distinction, certificate can be provided on request.

Membership

CryptoForma *2014 – dissolution*

Member of the CryptoForma network of excellence which aimed to build a UK-based network of security researchers in academia with particular focus on formal methods and cryptographic protocols.

AFiniteNumberOfMonkeys (University of Birmingham CTF Team) *2014 – 2015*

Undertook a lead role in organising the Capture The Flag team for the years 2014 and 2015. These years saw both the highest number of events participated in and highest ranking of the team to date.

SRSCC Member at University of Birmingham *2015 – 2016*

Member of committee for ensuring research students and staffs welfare, views and needs are identified and considered in administrative procedures within the School of Computer Science.

Relevant Skills

Programming languages (in order of proficiency):

`C`, `OCaml`, `Rust`, Assembly languages: `x86`, `x86-64`, `ARM`, `MIPS`, `Python`, `Erlang`, `C++`, `Haskell`.

Software & Frameworks:

`IDA Pro` (with `IDAPython`), `BAP` (`BinaryAnalysisPlatform`), `angr` (Concolic/DSE framework), `gdb` (with `PEDA`), `OllyDBG`, `WEKA` (machine learning framework).

Techniques & Technologies:

- Software exploitation (Linux & Web, to level of course instructor and CTF participant):
  - Binary exploitation: buffer (stack and heap) overflows, ROP (and variants), data-oriented attacks and associated mitigations.
  - Web-based exploitation: SQL injection, CSRF, XSS and associated mitigations.
- Reverse Engineering (Linux, Windows and raw `ARM/MIPS` based firmware, to level of defeating commercial software protection schemes, CTF participation, firmware reverse engineering, malware analysis).

Interpersonal:

Aptitude to work as part of a team: successfully collaborated on numerous deliverables for overseas client (Huawei); currently maintaining a paper collaboration with an academic at EURECOM (Aurélien Francillon) as well as working there in person for 3 weeks; published multi-authored academic papers; completed group programming project (achieving highest possible grade) – which considered teamwork in the grading process, as part of undergraduate degree.

Adaptability and ability to work under high pressure: took leading role in time-constrained, high pressure Capture The Flag competitions when faced with unknown challenges.

Transferring knowledge to others: undertaken teaching assistant responsibilities throughout Ph.D. and undergraduate degree; currently run hands-on lab session for Secure Programming course for around 50 students; provided weekly tutorials for University of Birmingham CTF team; designed and ran workshops for backdoor detection for client (Huawei); given numerous talks on research.

Spoken languages:

English *(native)*, Spanish *(elementary proficiency)*.

Outside Interests

In no particular order: art (Surrealism, in particular), philosophy (Baudrillard, Debord, Sartre), music (mostly Black Metal), films (Surrealist; David Lynch), clothes making, tea (drinking and collecting, especially Matcha and varieties of black tea).