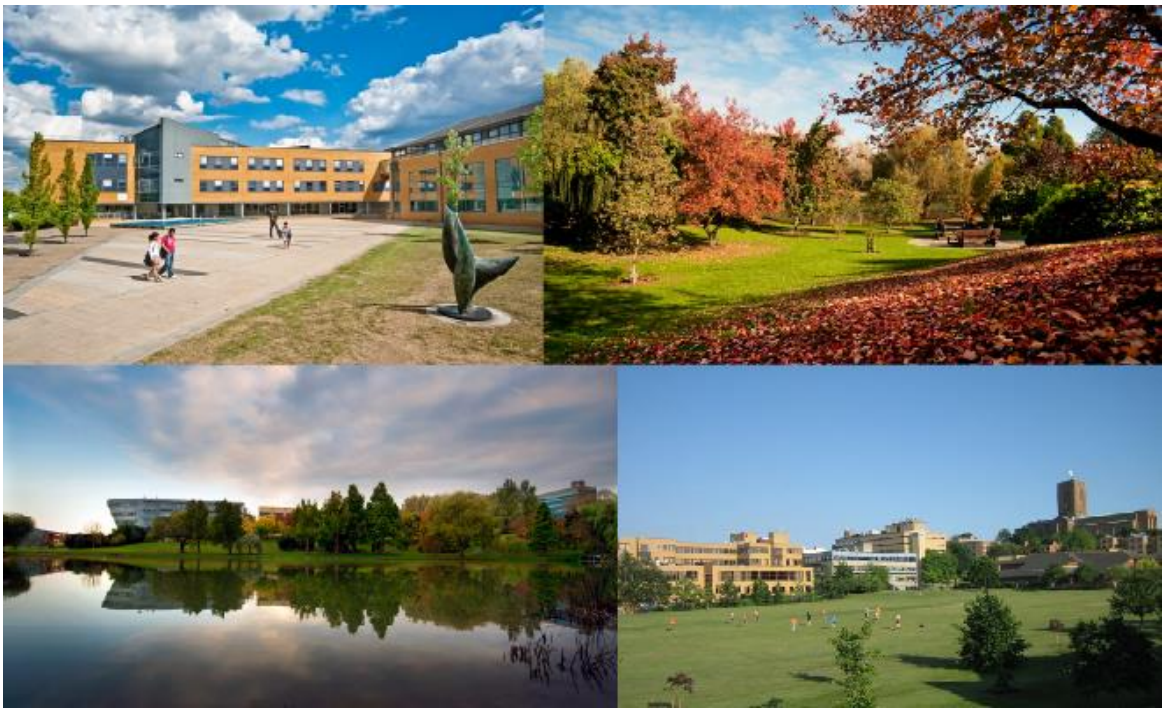
12th Annual Computing Department PhD Conference – CompConf 2015



Treetops, Wates House

University of Surrey

6th May 2015

Contents

Schedule of Events

Time	Event
Morning Session	
08:30	Registration (Tea and coffee will be served)
9:00	Opening Speech **Professor Steve Legg** (University Relations Head, IBM UK; Royal Academy of Engineering Visiting Professor, University of Surrey)
9:15	Keynote Speech 1 "Machine Learning in Computational Biology" **Professor Mahesan Niranjan** (University of Southampton)
10:00	**Oral Presentation Session 1** Chair: **Yin Hu** 10:00-10:20 **Shahrzad Shapoori** "A Novel Approach for Detection of Medial Temporal Discharges Using Blind Source Separation Incorporating Dictionary Look Up" 10:25-10:45 **Santosh Tirunagari** "Detection of Face Spoofing Using Visual Dynamics" 10:50-11:10 **Ran Cheng** "A Multiobjective Evolutionary Algorithm using Gaussian Process based Inverse Modeling"
11:15	Morning Tea and Coffee Break
11:45	**Oral Presentation Session 2** Chair: **Shahrzad Shapoori** 11:45-12:05 **Ataollah Ramezan Shirazi** "Collective Movement of a Swarm without Directional Sensing using Morphogen Signaling" 12:10-12:30 **Marcelo Damasceno** "Multi-Privacy Biometric Protection Scheme using Ensemble Systems" 12:35-12:55 **Santosh Tirunagari** "Detection and Evaluation of Finger Vein Spoofing in Monogenic Scale Space"
Buffet Lunch / Poster Session	
13:00	Poster Presentation Session **Accepted work: 1- Santosh Tirunagari, 2 - Shahrzad Shapoori** **Work in progress: 3 - Saeed Alqahtani, 4 - Andreas Antoniades, 5 - Hanan Alghamdi**
Afternoon Session	
14:15	Keynote Speech 2 "Cyber Security: What does the data tell us?" **Nick Coleman** (Global Head Cyber Security Intelligence, IBM)
15:00	**Oral Presentation Session 3** Chair: **Santosh Tirunagari** 15:00-15:20 **Nik Suki** "Deriving Confidence Intervals of Log-likelihood Ratio" 15:25-15:45 **Nouf Aljaffan** " Visual Password Checker (VPC): Making Proactive Password Checkers More Informative and Useful" 15:50-16:10 **Nurulhuda A Manaf** "Service Choreography, SBVR, and Time"
16:15	Afternoon Tea and Coffee Break / Judges Convene
16:40	Prize Presentation and Closing Speech **Dr Shujun Li** (PGR Director, Department of Computing)

Morning Session Keynote Speaker

Prof Mahesan Niranjan

School of Electronics and Computer Science
University of Southampton

Prof Niranjan joined the University of Southampton in February 2008 as a professor in the School of Electronics and Computer Science and served as Head of the Information: Signals, Images and Systems (ISIS) research group. Prior to joining Southampton, Prof Niranjan have worked for the University of Cambridge as a Lecturer in Information Engineering 1990-1998), and the University of Sheffield as Professor of Computer Science (1999-2007). At Sheffield he also served as Head of Computer Science (Feb 2002 - Aug 2004) and Dean of Engineering (Aug 2006 - Jan 2008). He is Chair of Staff Student Liaison Committee of the School of Electronics and Computer Science, University of Southampton.

Title of talk: Machine Learning in Computational Biology

Abstract: Machine learning is about extracting useful information from large and complex data-sets. Nowadays, with rapid growth in instrumentation, our ability to archive and distribute vast volumes of data and the willingness to share it, biology is becoming an exciting source of challenging machine learning problems. In this talk I will give a tutorial level introduction to the subject with particular focus on the regulation of cellular protein levels, embryonic development and systems level analysis of biological function.

Afternoon Session Keynote Speaker

Nick Coleman

Global Head Cyber Security Intelligence, IBM
Permanent Stakeholder, European Security Agency (ENISA)

Nick Coleman is Cyber Security Intelligence Services Head at IBM. Previously he was the UK Government Reviewer of Security and authored the 'Coleman Report'. He is an Honorary Professor at Lancaster University. He also serves on the EU Security Agency (ENISA) permanent stakeholders group. He is a founder of ExtraPaddle the mentoring platform which enables people to engage 3 hours a year to find / assist others with growth, knowledge and connections for their journeys.

Title of talk: Cyber Security: What does the data tell us?

Abstract: Cyber attacks dominate the headlines. This lecture will look at what we can understand from events, attacks and incidents, and the data behind them. We will look at whether the data we have can help us understand the nature of the problem? How the data we can collect and analyse today can help us understand the challenges we face. Then at the end the lecture will look to the future in terms of both the challenges and the opportunities that data will offer.

ORAL PRESENTATION ABSTRACTS

(Accepted/Published Work)

Session 1 - 10:00

A Novel Approach for Detection of Medial Temporal Discharges Using Blind Source Separation Incorporating Dictionary Look Up

Shahrzad Shapoori (with Saeid Sanei and Wenwu Wang)

In blind source separation (BSS), sparsity is proved to be very advantageous. If data is not sparse in its current domain, it can be modelled as sparse linear combinations of elements of a chosen dictionary. The choice of dictionary that sparsifies the data is very important. In this paper the dictionary is pre-specified based on chirplet modelling of various kinds of real epileptic spikes. Dictionary look up together with source separation is used to extract the closest source to the source of interest from the scalp EEG measurements. The algorithm has been tested on synthetic and real data consisting of epileptic discharges, and the results are compared with those of traditional BSS.

Session 1 - 10:25

Detection of Face Spoofing Using Visual Dynamics

Santosh Tirunagari

(with Norman Poh, David Windridge, Aamo Iorliam, Nik Suki and Anthony TS Ho)

Rendering a face recognition system robust is vital in order to safeguard it against spoof attacks carried out by using printed pictures of a victim (also known as print attack) or a replayed video of the person (replay attack). A key property in distinguishing a live, valid access from printed media or replayed videos is by exploiting the information dynamics of the video content, such as blinking eyes, moving lips, and facial dynamics. We advance the state of the art in facial anti-spoofing by applying a recently developed algorithm called Dynamic Mode Decomposition (DMD) as a general-purpose, entirely data-driven approach to capture the above liveness cues. We propose a classification pipeline consisting of DMD, Local Binary Patterns (LBP), and Support Vector Machines (SVM) with a histogram intersection kernel. A unique property of DMD is its ability to conveniently represent the temporal information of the entire video as a single image with the same dimensions as those images contained in the video. The pipeline of DMD+LBP+SVM proves to be efficient, convenient to use, and effective. In fact only the spatial configuration for LBP needs to be tuned. The effectiveness of the methodology was demonstrated using three publicly available databases: print-attack, replay-attack, and CASIA-FASD, attaining comparable results with the state of the art, following the respective published experimental protocols.

Session 1 - 10:50

A Multiobjective Evolutionary Algorithm using Gaussian Process based Inverse Modeling

Ran Cheng (with Yaochu Jin, Kaname Narukawa and Bernhard Sendhoff)

To approximate the Pareto front, most existing multiobjective evolutionary algorithms store the non-dominated solutions found so far in the population or in an external archive during the search. Such algorithms often require a high degree of diversity of the stored solutions and only a limited number of solutions can be achieved. By contrast, model-based algorithms can alleviate the requirement on solution diversity and in principle, as many solutions as needed can be generated. This paper proposes a new model-based method for representing and searching non-dominated solutions. The main idea is to construct Gaussian process based inverse models that map all found non-dominated solutions from the objective space to the decision space. These inverse models are then used to create offspring by sampling the objective space. To facilitate inverse modeling, the multivariate inverse function is decomposed into a group of univariate functions, where the number of inverse models is reduced using a random grouping technique. Extensive empirical simulations demonstrate that the proposed algorithm exhibits robust search performance on a variety of medium to high dimensional multiobjective optimization test problems. Additional non-dominated solutions are generated a posteriori using the constructed models to increase the density of solutions in the preferred regions at a low computational cost.

Session 2 - 11:45

Collective Movement of a Swarm without Directional Sensing using Morphogen Signaling

Ataollah Ramezan Shirazi (with Yaochu Jin)

In this paper, we present a morphogenetic approach to self-organized collective movement of a swarm. We assume that the robots (agents) do not have global knowledge of the environment and can communicate only locally with other robots. In addition, we assume that the robots are not able to perform directional sensing. To self-organize such systems, we adopt here a simplified diffusion mechanism inspired from biological morphogenesis. A guidance mechanism is proposed based on the history of morphogen concentrations. The division of labor is achieved by type differentiation to allocate different tasks to different type of robots. Simulations are run to show the efficiency of the proposed algorithm. The robustness of the algorithm is demonstrated by introducing an obstacle into the environment and removing a subset of robots from the swarm.

Session 2 - 12:10

Multi-Privacy Biometric Protection Scheme using Ensemble Systems

Marcelo Damasceno (with A.M.P Canuto and Norman Poh)

Biometric systems use personal biological or behavioural traits that can uniquely characterise an individual but this uniqueness property also becomes its potential weakness when the template characterising a biometric trait is stolen or compromised. To this end, we consider two strategies to improving biometric template protection and performance, namely, (1) using multiple privacy schemes and (2) using multiple matching algorithms. While multiple privacy schemes can improve the security of a biometric system by protecting its template; using multiple matching algorithms or similarly, multiple biometric traits along with their respective matching algorithms, can improve the system performance due to reduced intra-class variability. The above two strategies lead to a novel, ensemble system that is derived from multiple privacy schemes. Our findings suggest that, under the worst-case scenario evaluation where the key or keys protecting the template are stolen, multi-privacy protection scheme can outperform a single protection scheme as well as the baseline biometric system without template protection.

ORAL PRESENTATION ABSTRACTS

(Work in Progress/Under Review)

Session 2 - 12:35

Detection and Evaluation of Finger Vein Spoofing in Monogenic Scale Space

Santosh Tirunagari (with Shekar Hema B, Norman Poh, Saeid Sanei and Anthony TS Ho)

Recent studies have shown that it is possible to attack a biometric system that is based on finger vein by using printed materials. In this study, we propose to detect spoofing in finger vein using monogenic scale space (MSS) based global descriptor that represents an image by using a distribution of orientation information weighted by its local energy. The proposed method compares favourably with established methods such as discrete cosine transforms and entropy based filter methods. In all experiments, SVM with a histogram intersection kernel is used. Our two-fold subjects-based cross-validation experiments show that the printed finger vein images contain artifacts, have different image quality, and light reflection property than valid/live finger vein images.

Session 3 - 15:00

Deriving Confidence Intervals of Log-likelihood Ratio

Nik Suki (with Norman Poh)

Forensic identification is the task of determining whether or not observed evidence coming from a known source. It involves determining a log likelihood ratio (LLR) given a piece of evidence (E) under the prosecution hypothesis (H0, the evidence came from the same source) and the defence hypothesis (H1, the evidence did not came from the same source). In LLR(E) based decision methods, a variable number of input evidences are used. A decision based on many pieces of evidence can result in nearly the same LLR(E) as one based on few pieces of evidence. In this paper, we will evaluate the LLR(E) performance by measuring its calibration. Later we compare the confidence intervals of the calibrated LLR(E) versus normal LLR(E) using the traditional bootstrapping method. At the end of this study, we will show how the proposed calibration may affect the system performance of LLR(E) by using face and speech biometric.

Session 3 - 15:25

Visual Password Checker (VPC): Making Proactive Password Checkers More Informative and Useful

Nouf Aljaffan (with Shujun Li and Kyriakos Kafas)

A proactive password checker is a software component showing besides a password input box for displaying the estimated password strength, which is designed to help human users to define stronger passwords and sometimes used to inform users about the established

password policy as well. Although evidence has showed that proactive password checkers can indeed help users to define stronger passwords, all existing proactive password checkers are visualised using very simple approaches such as one or more one-dimensional colour bars and/or textual descriptions about the strength of the password being checked. Some proactive password checkers can also mislead users if the password metering algorithm used is not accurate. In this work, we propose to extend the concept of proactive password checkers to a *multi-dimensional visualiser* of *more* information about password security, including password strengths estimated by different password metering algorithms and threats related to different types of weak passwords and attacks to passwords, in order to better inform users about the complexity of password strength estimation and the existence of diverse attacks to passwords so that they can make more informed decisions about passwords for different application scenarios. In other words, we propose to enrich the way of how visualisation is done for proactive password checkers to make them more informative and useful (and less misleading). In this work we call this new password security visualiser *Visual Password Checker* (VPC). We designed and implemented a web-based prototype of VPC in 2-D space for textual passwords and we are in the process of running some lab- and crowdsourcing-based user studies to evaluate the performance of the implemented VPC against 3 other types of existing proactive password checkers, and it is hypothesised that the implemented VPC will be able to educate users more about password security and can persuade them to define stronger passwords. Future work includes more visualisation items, extension to 3-D space, enhanced user interface, and support for graphical passwords.

Session 3 - 15:50

Service Choreography, SBVR, and Time

Nurulhuda A Manaf (with Sotiris Moschoyiannis)

Services are the basic elements used in the Service Oriented Computing (SOC) paradigm to develop distributed applications. However, the coordination of the underlying interactions between loosely-coupled services, across different organisations over the web, remains an open challenge. The service choreography approach has been exploited in setting up the interactions between services since it captures the main aspect in the interactions; that is, the orderings of invocations on service interfaces. In this paper, we propose the use of SBVR as a modelling language for service choreography, which focuses on what rather than how the choreography should be played out. The declarative approach we propose is extended with a notion of time and this allows to capture the ordering of service interactions, which is crucial in the coordination. The expression of time ordering constraints in a choreography specification is done in a way that respects the vocabulary and business rules of SBVR, as advocated by OMG, without adding special primitives or changing the specification of SBVR itself.

POSTER PRESENTATION ABSTRACTS

Note: Posters 1 and 2 have been accepted for publication, so neither will participate in the competition for the best poster prize which is mainly for posters describing early research work/plans of Year-1 PGR students.

1

Automatic Classification of Long-term Kidney Functions Using Machine Learning Techniques term Kidney Functions Using Machine Learning Techniques

Santosh Tirunagari (with Norman Poh, Simon de Lusignan and Farmer Christopher)

Screening through thousands of chronic kidney disease (CKD) patients for identifying those who are at risk is extremely important, and no automated methods exist to solve this problem. This is the first reported attempt that compares automated methods with human annotation. We consider eGFR time-series of 488 patients with CKD at various stages. For each eGFR time-series, 5 consultants annotated as being stable, linear or step-change. The class label assigned to an eGFR time-series is based on the consensus of these five annotations. In order to develop an automated approach to classify, the time-series is first modelled using Gaussian process regression (GPR) and resampled using a fixed-size vector of 50 observations in time. A supervised machine learning technique called K-Nearest Neighbour (K-NN) classifier is then used to classify the trends through a 5-fold cross-validation experimental protocol. We found that the classifier employed automatically classified the eGFR trend with an accuracy of 38.5% for the linear trend (of eGFR), 85.7% for the stable trend, and 62.9% for the step-change trend. In comparison, a human expert can do so at 92.3%, 98.4%, and 93.7% accuracy, respectively.

2

A Novel Approach for Detection of Medial Temporal Discharges Using Blind Source Separation Incorporating Dictionary Look Up

Shahrzad Shapoori (with Saeid Sanei and Wenwu Wang)

In blind source separation (BSS), sparsity is proved to be very advantageous. If data is not sparse in its current domain, it can be modelled as sparse linear combinations of elements of a chosen dictionary. The choice of dictionary that sparsifies the data is very important. In the proposed method, the dictionary is pre-specified based on chirplet modelling of various kinds of real epileptic spikes. In this work, a hybrid dictionary look up – source separation is introduced to extract the closest source to the source of interest from the scalp EEG measurements. The algorithm has been tested on synthetic and real data consisting of epileptic discharges, and the results are compared with those of traditional BSS.

3

Mobile Privacy Leakage Detection and Prevention

Saeed Alqahtani (with Shujun Li and Anthony TS Ho)

Mobile devices are ubiquitous in today's digital world. While people enjoy the convenience brought by mobile devices, it has been found that many mobile apps leak personal information without user awareness. That can occur due to careless programming errors and/or intention of app developers to collect private information for various reasons. To better protect users, the research community has made a lot of effort to design better systems for detecting and preventing mobile privacy leakages. However, most existing tools simply produce notifications for detected leakages and leave further actions to end users who often lack experience and/or information to make a proper decision and annoyed by false alerts. Therefore, one goal of this PhD project is to investigate how users can be better engaged to increase accuracy of privacy leakage detection while minimising unnecessary user intervention. Moreover, most those tools use a static list of information sources covering limited privacy leakages. Another aim of this project is thus to cover more dynamic and user-specific information sources neglected by existing tools. A third aim is to add an active forensic element so that more information about how leaked personal information is used can be obtained. To support the above aims, we are currently developing a privacy leakage protection benchmarking system which will be used to analyse a large number of existing tools in the market and developed by researchers.

4

A Sample-and-Learn scheme for Neural Networks

Andreas Antoniades (with Clive Cheong Took)

Designing scalable neural networks for big data is a formidable task. The common solution for the problem of big data is to employ distributed computing where each thread must be independent from each other. Yet, the inherent structure of a neural network makes it difficult to adapt for parallel computing due to the inter-dependency between neurons. Moreover, the bigger the neural network is, the more parameters need to be optimised. To circumvent those issues, we propose a simple yet effective way (i.e. sample-and-learn scheme) to render neural networks scalable to big data by sampling a much smaller subset of the data at hand. It is the first time that reservoir sampling is introduced for the training of neural network learning. As an application, we consider the large-scale problem of forecasting hourly electric power consumption for a period of approximately four years based on a dataset that includes measurements such as global active power, global reactive power, voltage and global intensity. Another issue with the big dataset is its veracity - missing samples in measurements, which is overcome naturally by our sampling solution. Simulation results on this real-world dataset support our approach, achieving an accuracy of more than 95% even when less than 1% of the dataset is used for the learning of the neural network.

5

Retinal Fundus Image Analysis Project

Hanan Alghamdi (with Su Wang and Lilian Tang)

This work presents the current and future works of the Retinal Fundus Image Analysis Project. In this project a computer aided diagnosis system is to be developed to aid in the diagnosis of several eye diseases such as glaucoma and diabetic retinopathy. Early detection of pathological signs is critical to prevent disease progression and vision loss. Normal structure of the retina, which includes the optic disc and blood vessels, should be extracted before the detection of any clinical signs. Several image processing and machine learning methods have been evaluated to detect optic disc, blood vessels, dark and bright lesions. Two approaches have been utilized for optic disc detection based on two assumptions: the geometrical relationship between optic disc and the blood vessels and the colour, shape, size and vessel density of the optic disc. The results show the advantage of both approaches. For blood vessel segmentation line tracking and ensemble learning were utilized. The tracking approach is mainly based on the curve fragments generated with vessel tracking. This overcomes the drawbacks in segmenting false vessels around the optic disc and is able to extract fine or tortuous vessels. In ensemble learning approach, several statistical diversity measures were utilized to optimize ensembles used to segment blood vessels. Ant Colony Optimization (ACO) is adopted to select the ensemble base models using various criteria. The results demonstrate the necessity and the advantage of ensemble optimization to support the blood vessel segmentation. Microaneurysms, which appear as dark lesions were detected by adopting singular spectrum analysis on cross sectional profiles of candidate regions and the correlation coefficient between each profile and a typical microaneurysm profile. The results show the robustness of the approach when testing on large scale datasets with clinically acceptable sensitivity and specificity.

Organising Committee

- Ran Cheng (PhD Rep)
- Aamo Iorliam (PhD Rep; Local Arrangement Chair)
- Veronika Kuchta (PhD Rep; General Co-Chair)
- Dr Shujun Li (PGR Director; General Co-Chair)
- Denise Myers (Departmental Secretary)
- Shahrzad Shapoori (PhD Rep; Publicity Chair)
- Santosh Tirunagari (PhD Rep)
- Sarah Turnbull (Research Administrative Assistant; Finance Chair)

Technical Programme Committee

- Ran Cheng (PhD Rep; Co-Chair)
- Yin Hu (PhD Student)
- Aamo Iorliam (PhD Rep)
- Veronika Kuchta (PhD Rep)
- Dr Shujun Li (PGR Director; Co-Chair)
- Shahrzad Shapoori (PhD Rep)
- Santosh Tirunagari (PhD Rep)

Prize Judging Panel

- Prof Gloria Benson – Consult Hyperion (Director) and University of Surrey (Visiting Professor)
- Prof Anthony TS Ho – University of Surrey (Head of Multimedia, Security and Forensics Research Group)
- Prof Steve Legg (Chair) – IBM UK (University Relations Head) and University of Surrey (Royal Academy of Engineering Visiting Professor)
- Matt Lewis – NCC Group (Executive Principal Consultant)
- Dr Mark Manulis – University of Surrey (Deputy Director, Surrey Centre for Cyber Security)
- Tony Phipps – Lloyds Banking Group (Senior Manage, Digital Fraud & Security)
- Dr Loukianos Spyrou – University of Surrey (Postdoctoral Research Fellow)
- Prof Adrian Waller – Thales UK Ltd (Senior Security Consultant) and University of Surrey (Visiting Professor)

Reviewers

- Dr Anna Lisa Ferrara (Academic) – Department of Computing, University of Surrey
- Dr Lee Gillam – Department of Computing, University of Surrey
- Patrik Huber (PhD Student) – Centre for Vision, Speech and Signal Processing (CVSSP), Department of Electronic Engineering, University of Surrey
- Franziskus Kiefer (PhD Student) – Department of Computing, University of Surrey
- Prof Paul Krause (Academic) – Department of Computing, University of Surrey
- Veronika Kuchta (PhD Student) – Department of Computing, University of Surrey
- Prof Yaochu Jin (Academic) – Department of Computing, University of Surrey
- Dr Shujun Li (Academic) – Department of Computing, University of Surrey
- Dr Mark Manulis (Academic) – Department of Computing, University of Surrey
- Dr Saeid Sanei (Academic) – Department of Computing, University of Surrey
- Santosh Tirunagari (PhD Student) – Department of Computing, University of Surrey
- Dr Hongying (Lillian) Tang (Academic) – Department of Computing, University of Surrey
- Dr Stephan Wesemeyer (Academic) – Department of Computing, University of Surrey

Photographers

- Marcelo Damasceno – Department of Computing, University of Surrey (Visiting PhD Student) / Federal University of Rio Grande do Norte, Brazil (PhD Student) / Federal Institute of Rio Grande do Norte, Brazil (Lecturer)
- Ataollah Ramezan Shirazi – Department of Computing, University of Surrey (PhD Student)
- Andreas Theodorou – Department of Computing, University of Surrey (Final-Year UG Student)

Sponsors


IBM


nccgroup
freedom from doubt


LLOYDS BANK


consult
hyperion
securing
tomorrow's
transactions