Research Statement

Hamed Haddadi

January 2022

My research focus is on networked and user-centred sensing systems. My primary objective is to design, build, implement, and evaluate networked systems which allow us to efficiently collect, analyse, and utilise the rich sources of personal data available to us, without jeopardising our individual privacy and security. Research in this area entails dealing with the increasing number of devices and services surrounding us (e.g., Internet of Things devices), and those *on* us (e.g., smartphones and wearables), alongside the Machine Learning (ML) models that run on these devices, leveraging the growing speed, memory, and sensing capabilities of these cyber-physical systems.

Research philosophy

A foundation objective of my research is to enable the individuals to interact with, and act upon, the inferences drawn from their data for their personal insight, or the societal good (e.g., public health and wellbeing applications building on aggregation of personal data from anonymous individuals). My research into CPS is focused primarily on the practical aspects of privacy, security, machine learning, and systems and networks challenges. My approach to research is largely interdisciplinary. I rely on harnessing theoretical foundations from different disciplines, then extensive measurements and user studies to understand the challenges and contain the practical and real problems. I then work on design and implementation of systems and algorithms to provide direct solutions to these challenges, and then evaluate these in the wild.

My research approach is on building and evaluating real, practical software and hardware systems which individuals interact with. A pivotal aspect in my work is the combination of large-scale personal and public data collection from mobile sensors and social media, practical machine-learning algorithms for analyses and inferences on these data to turning them into useful information. The information can then be used for designing sensing and analytics systems with an emphasis on respect for individuals' privacy, security, and resources. I explore methods of performing analytics and correlations on the *edge* (e.g., smartphone apps, personal data hubs, and ambient sensors) to draw inferences from missing or partial data towards achieving beneficial utility and practical privacy. My research often leads to collaborations with researchers across a wide spectrum, from psychologists and physicians, to engineers and ethnographers. I enjoy this knowledge exchange.

RESEARCH CONTRIBUTIONS

My research has been interdisciplinary, with a focus on user-centred system. I have had a number of significant contributions in this space; I will briefly describe a few here.

Privacy-Preserving Systems

I have been working in the space of privacy-preserving sensing and analytics for a decade now, starting with my earlier works on design of behavioural analytics, advertising, and fraud systems (*ACM HotNets 2019, CCR 2010, MobiArch 2010, ACM IMC 2013*). My works in privacy-aware analytics and fraud detection have been adopted by corporations such as Microsoft, and Brave[1] (where I Currently serve as the Chief Scientist and the Visiting Professor). More recently, I have been working on methods to leverage the increasingly advanced processing and memory capabilities at the edge, i.e., smartphones and personal devices, to provide hybrid platforms for high-utility, yet private, analytics platform. These include models relying on splitting deep learning models (*IEEE TKDE 2019, IEEE ISIT 2019,*

[1] https://brave.com/

IEEE IoTJ 2020), or design of personalised models to be executed mostly on the client side (*ACM/IEEE IoTDI 2018 & 2019, ACM MobiSys 2020, MLSys 2020, Usenix NSDI 2022*).

Personal Data Sensing & Analytics
In the last few years, I led the Databox[2] project, an open-source platform privacy-aware IoT and personal data analytics. As the project PI, I led the research in data analytics, while leveraging on my works in understanding user behaviours through data from wearables, smartphones, and social media (*ICWSM 2009 & 2010, ACM CCR 2012, PETS 2015*). Databox was built on the principles of Human-Data Interaction (HDI)[3], a research community that I co-created and lead since 2013 in conjunction with industry partners such as BBC R&D, Google, and Telefonica Research. HDI is a framework centred on providing legibility, negotiability, and agency on inference and use of personal data (*DE 2013, PUC 2016, IET Book 2021*). I am currently a co-I on the EPSRC Human-Data Interaction NetworkPlus.

Recently, in the EPSRC DADA project, my group has been looking at information exposure from IoT devices. We run one of the largest IoT testbeds in the world,[4] where we look at a number of privacy and security mechanisms for IoT devices (ACM IMC 2019, ACM IMC 2020, PETS 2020). Our opensource IoT privacy and security platform, IoTrim (PETS 2021),[5] has received a recent a generous cash award from Deutsche Telekom, and we are considering commercial opportunities in this space. Our work in this space has been part of several documentaries (BBC, NYTimes, Channel 4, Which?) and a few government investigations (CDEI, CMA, ICO).

Wearables and Mobile Sensing
In the past decade I have been leading several research efforts in the space of sensing. This started form my postdoctoral research period at the Royal Veterinary College and the Department of Pharmacology at University of Cambridge, where I implemented the sensing and data analysis platform for the study of a human model of the Huntington Disease using Merino sheep (*Behavioral Ecology and Sociobiology 2011, Current Biology 2012*). Following this, I worked for a while on smartphone sensing and social media data analysis for investigating childhood obesity patterns (*IEEE ICHI 2015, ACM Digital Health 2015, ICWSM 2016, WWW 2017*). I have also worked on exploring the usability of wearables for individuals' physical and psychological wellbeing (*ACM ISWC 2015, ACM CHI 2018*). These works have also led to SensingKit[6] and AWSense, popular smartphone and wearable Sensing libraries. Recently, I have been working on privacy-preserving analytics models on these devices, as part of my work as a Co-I on the £20m UK Dementia Research institute at Imperial College London (ACM Mobisys 2021 Best Paper, IEEE PerCom 2021, ACM IMWUT/UBICOMP 2021).

Many of my research publications have been presented at leading media including the BBC, NYTimes, Washington Post, Harvard Business Review (see my webpage for details), and have led to a number of large grants and new initiatives including DARPA and EPSRC calls on Human Data Interaction, and the EPSRC NetworkPlus on Human-Data Interaction.

CURRENT & FUTURE RESEARCH

My research around User-Centred Cyber-Physical Systems in the next few years will focus on:

Securing the Next Billion Consumer Devices on the Edge: Beyond the original vision of the EPSRC Databox project, I constantly look for new ways of utilising the power of the modern cyber-physical systems at the edge to augment the cloud. In this space, I am mow

[2] https://github.com/me-box/databox
[3] http://hdiresearch.org
[4] https://www.imperial.ac.uk/systems-algorithms-design-lab/research/advanced-iot-testbed/
[5] https://iotrim.github.io/
[6] https://www.sensingkit.org

looking at novel ways of utilising data form the range of IoT devices available in individuals' household, to their shopping activities, and social media data, to provide interesting, yet private and personalised applications for the users. One of the approaches I am currently investigating is the use of Trusted Execution Environments in edge devices (e.g., ARM TrustZone in smartphones) for personalised and federated learning applications (ACM MobiSys 2020 and 2021 Best Paper Award). In this way, one can split larger models into layers that can be executed on the TEE and the cloud in a hybrid manner, while optimising various constraints (e.g., privacy, utility, and energy). In the next fives years, as part of an EPSRC OpenPlus Fellowship, I will be pursuing this research as part of a major collaborations with ARM, Telefonica, CISCO, and Samsung AI, alongside efforts in the regulator landscape with the UK Information Commissioner's Office.

In related research, jointly with Brave, I am looking at the use of the large variety of signals and capabilities in client devices, to provide private performance analytics, model attestation, and reputation metrics for client authenticity. These are important specifically in the context of client-side content personalisation, and fraud detection.

Understanding CPS applications *in-the-wild*: In my lab, we have a comprehensive test-bed of IoT devices, equipped with advanced data collection infrastructure, access to the latest devices including high-frequency power analysis platforms (ACM HotNets 2019, PAM 2022). I am establishing a "home-like" environment in the lab with furniture and appliances, based on a number of exhibitions with the BBC and the Databox team as the "Living Room of the Future". While these environments are great for establishing baseline behaviour of devices and systems, I have a passion for understanding the evaluation of theoretical and optimal systems and devices in real-life scenarios with real users.

In the next 5 years, I will continue to work on applications which will enhance the daily lives of dementia patients, while ensuring their privacy and independence. These applications include behaviour analysis, daily routine monitoring, indoor and outdoor navigation and localisation, and feedback and interventions. Performing experiments with such user groups is a highly delicate and challenging tasks, involving physicians, technologist, ethnographers, and psychologists and I will leverage my interdisciplinary experience for delivering the objectives of this project.

Future Agenda

My longer term research vision is directed towards building a user-centred CPS ecosystem which is privacy-aware, secure, efficient, and reliable. I will expand my research efforts, collaborations, and industry engagement. I aim to do more out of the lab, with real deployments and evaluations, leveraging my experience in carrying out real world user studies. I will perform concrete evaluations of the cyber-physical systems that I am working on, from bare silicon resource utilisation, to algorithmic privacy and usability bounds, all the way to user experience evaluation. I look for interesting and exciting, yet fundamental and long-term research problems. Sometimes it takes years to establish the right links and connect the disciplines for solving these challenges. It can be burdensome to measure the depth of the problem and the characteristics of those affected by it, to design solutions and systems to overcome the obstacles, and to understand the real-life constraints around the problem space. However, the fruit of such research investments is certainly more valuable, and this approach will be the underpinning of my research in the years to come alongside publications and commercialisation opportunities arising from the research.