DAGSTUHL
REPORTS

Volume 6, Issue 9, September 2016

ISSN 2192-5283

Published online and open access by
Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at http://www.dagstuhl.de/dagpub/2192-5283

Publication date
February, 2017

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at `http://dnb.d-nb.de`.

Digital Object Identifier: 10.4230/DagRep.6.9.i

Aims and Scope
The periodical *Dagstuhl Reports* documents the program and the results of Dagstuhl Seminars and Dagstuhl Perspectives Workshops.
In principal, for each Dagstuhl Seminar or Dagstuhl Perspectives Workshop a report is published that contains the following:

- an executive summary of the seminar program and the fundamental results,
- an overview of the talks given during the seminar (summarized as talk abstracts), and
- summaries from working groups (if applicable).

This basic framework can be extended by suitable contributions that are related to the program of the seminar, e. g. summaries from panel discussions or open problem sessions.

Contact
Schloss Dagstuhl – Leibniz-Zentrum für Informatik
Dagstuhl Reports, Editorial Office
Oktavie-Allee, 66687 Wadern, Germany
reports@dagstuhl.de
http://www.dagstuhl.de/dagrep

Report from Dagstuhl Seminar 16361

Network Attack Detection and Defense – Security Challenges and Opportunities of Software-Defined Networking

Edited by

Marc C. Dacier[1], Sven Dietrich[2], Frank Kargl[3], and Hartmut König[4]

1 **QCRI – Doha, QA,** mdacier@qf.org.qa
2 **City University of New York, US,** spock@ieee.org
3 **Universität Ulm, DE,** frank.kargl@uni-ulm.de
4 **BTU Cottbus, DE,** hartmut.koenig@b-tu.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 16361 "Network Attack Detection and Defense: Security Challenges and Opportunities of Software-Defined Networking".

Software-defined networking (SDN) has attracted a great attention both in industry and academia since the beginning of the decade. This attention keeps undiminished. Security-related aspects of software-defined networking have only been considered more recently. Opinions differ widely. The main objective of the seminar was to discuss the various contrary facets of SDN security. The seminar continued the series of Dagstuhl events Network Attack Detection and Defense held in 2008, 2012, and 2014. The objectives of the seminar were threefold, namely (1) to discuss the security challenges of SDN, (2) to debate strategies to monitor and protect SDN-enabled networks, and (3) to propose methods and strategies to leverage on the flexibility brought by SDN for designing new security mechanisms. At the seminar, which brought together participants from academia and industry, we discussed the advantages and disadvantages of using software-defined networks from the security point of view. We agreed that SDN provides new possibilities to better secure networks, but also offers a number of serious security problems which require further research. The outcome of these discussions and the proposed research directions are presented in this report.

Seminar September 4–9, 2016 – http://www.dagstuhl.de/16361
1998 ACM Subject Classification C.2.1 Network Architecture and Design, C.2.3 Network Operations, K.6.5 Security and Protection
Keywords and phrases attack detection, denial-of-service attack detection and response, intrusion detection, malware assessment, network monitoring, openflow protocol, programmable networks, security, software-defined networking, targeted attacks, vulnerability analysis
Digital Object Identifier 10.4230/DagRep.6.9.1
Edited in cooperation with Radoslaw Cwalinski

Network Attack Detection and Defense – Security Challenges and Opportunities of Software-Defined Networking, *Dagstuhl Reports*, Vol. 6, Issue 9, pp. 1–28
Editors: Marc C. Dacier, Sven Dietrich, Frank Kargl, and Hartmut König

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Hartmut König
Marc C. Dacier
Sven Dietrich
Frank Kargl
Radoslaw Cwalinski

From September 4 through 9, 2016, more than 40 researchers from the domains of computer networks and cyber security met at Schloss Dagstuhl to discuss security challenges and opportunities of software-defined networking (SDN).

Software-defined networking has attracted a great attention both in industry and academia since the beginning of the decade. This attention keeps undiminished. In 2014, IDC predicted that the market for SDN network applications would reach $1.1bn. Especially in industry, the vision of "programming computer networks" has electrified many IT managers and decision makers. There are great expectations regarding the promises of SDN. Leading IT companies, such as Alcatel-Lucent, Cisco systems, Dell, Juniper Networks, IBM, and VMware, have developed their own SDN strategies. Major switch vendors already offer SDN-enabled switches.

Software-defined networking provides a way to virtualize the network infrastructure to make it simpler to configure and manage. It separates the control plane in routers and switches, which decides where packets are sent, from the data plane, which forwards traffic to its destination, with the aim to control network flows from a centralized control application, running on a physical or virtual machine. From this controller, admins can write and rewrite rules for how network traffic, data packets, and frames are handled and routed by the network infrastructure. Routers and switches in a sense become "slaves" of this application-driven central server. SDN-enabled networks are capable of supporting user requirements from various business applications (SLAs, QoS, Policy Management, etc.). This is not limited to the network devices of a certain vendor. It can be applied to devices from various vendors if the same protocol is used. Most SDN infrastructure utilizes the widely-used OpenFlow protocol and architecture to provide communication between controllers and networking equipment.

Security-related aspects of software-defined networking have only been considered more recently. Opinions differ widely. Some believe that the security problems introduced by SDN are manageable – that SDN can even bring security benefits; others think that Pandora's Box has been opened where SDN and SDN-enabled networks can never be secured properly.

No doubt, there are a number of serious security problems as the following examples show. SDN controllers represent single points of failures. The controllers as well as the connections between controllers and network devices might be subject to distributed denial of service attacks. Compromising the central control could give an attacker command of the entire network. The SDN controllers are configured by network operators. Configuration errors can have more complex consequences than in traditional settings because they may unpredictably influence the physical network infrastructure. Furthermore, the idea of introducing 'network applications' that interact with the controller to modify network behavior seems like a complexity nightmare in terms of required authentication and authorization schemes. Finally, the SDN paradigm is a major turn around with respect to the basic design rules that have made the Internet successful so far, namely a well-defined layered approach. Whereas in

today's world, applications have no say in routing decisions, SDN's promise for highly flexible and application-tailored networking requires a way for applications to optimize networking decisions for their own benefits. However, it is unclear to what extent fairness can be ensured, how conflicting decisions can be resolved, etc. Along the same line, members of the security community worry about the possibility to intentionally design SDN applications that could eventually be turned into attack weapons or simply be misused by malicious attackers. Whether these fears are substantiated or not is something which has not received any scrutiny so far.

On the other hand, SDN is also considered by many researchers as an effective means to improve the security of networks. SDN controllers can be used, for instance, to store rules about the permission of certain requests which cannot be decided at the level of a single switch or router because this requires full overview over network status or additional information and interactions which are not contained in the current protocol versions. Attacks that can be detected this way are ARP spoofing, MAC flooding, rogue DHCP server, and spanning tree attacks. Also, by enabling the creation of virtual networks per application, people speculate that intrusion detection techniques relying on the modeling of the normal behavior of network traffic will become much easier to implement and more reliable in terms of false positive and negatives. Similarly, SDN apps could offer a very simple and effective way to implement quarantine zones for infected machines without cutting them off completely from the network since the quarantine could be customized at the application level (letting DNS and HTTP traffic for a given machine go through but not SMTP, for instance).

These two contrary facets of SDN security were the key ingredients for an extremely lively and very fruitful seminar. The seminar brought together junior and senior experts from both industry and academia, covering different areas of computer networking and IT security. The seminar started with two invited talks by Boris Koldehofe (TU Darmstadt, DE) and Paulo Jorge Esteves-Veríssimo (University of Luxembourg, LU) on the basics and security aspects of software-defined networking. After that we organized six working groups to discuss in two rounds the Good and the Bad of using SDN from the security point of view. Based on the outcome of the working groups and a plenary discussion, we formed another four working groups to discuss required research directions. The first six working groups focus on the following issues: (1) centralization in SDN, (2) standardization and transparency, (3) flexibility and adaptability for attackers and defenders, (4) complexity of SDN, (5) attack surface and defense, and (6) novelty and practicability. The research direction working groups dealt with (1) improving SDN network security, (2) a secure architecture for SDN, (3) secure operation in SDN-based environments, and (4) SDN-based security. The discussion in the working groups was supplemented by short talks of participants to express their positions on the topic or to report about ongoing research activities. Based on the talks, discussions, and working groups, the Dagstuhl seminar was closed with a final plenary discussion which summarized again the results from the working groups and led to a compilation of a list of statements regarding the security challenges and opportunities of software-defined networking. The participants agreed that SDN provides new possibilities to better secure networks, but also offers a number of serious security problems which have to be solved for being SDN a successful technology. The outcome of these discussions and the proposed research directions are presented in the following.

2 Table of Contents

3 Invited Talks

3.1 An overview on Software-defined Networking

Boris Koldehofe (TU Darmstadt, DE)

Joint work of Frank Dürr, Boris Koldehofe

Software-defined networking is currently a big trend in networking with strong support from both academia and industry. The basic concept of SDN is the separation of network control (control plane) and forwarding functionality (forwarding plane). The control plane is implemented by a controller hosted on a server, which programs the forwarding tables of switches to define communication "flows" in the network. Formerly distributed control logic like distributed routing algorithms are replaced by logically centralized control based on a global view onto the network. This talk discusses the motivation of SDN, offers a basic introduction of the corresponding concepts, and discusses some fundamental challenges.

3.2 Towards Secure and Dependable Software-Defined Networks

Paulo Jorge Esteves-Veríssimo (University of Luxembourg, LU)

Joint work of Paulo Jorge Esteves-Veríssimo, Diego Kreutz, Fernando Ramos
Main reference D. Kreutz, F. M. V. Ramos, P. Verissimo, "Towards secure and dependable software-defined networks", in Proc. of the 2nd ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN'13), pp. 55–60, ACM, 2013.
URL http://dx.doi.org/10.1145/2491185.2491199

Software-defined networking empowers network operators with more flexibility to program their networks. With SDN, network management moves from codifying functionality in terms of low-level device configurations to building software that facilitates network management and debugging. By separating the complexity of state distribution from network specification, SDN provides new ways to solve long-standing problems in networking, e.g., routing, while simultaneously allowing the use of security and dependability techniques, such as access control or multi-path. However, the security and dependability of the SDN itself is still an open issue. In this position paper we argue for the need to build secure and dependable SDNs by design. As a first step in this direction, we describe several threat vectors that may enable the exploit of SDN vulnerabilities. We then sketch the design of a secure and dependable SDN control platform as a materialization of the concept advocated here. We hope that this paper will trigger discussions in the SDN community round these issues and serve as a catalyser to join efforts from the networking and security & dependability communities in the ultimate goal of building resilient control planes.

4 Overview of Talks

4.1 Network Monitoring & SDN

Johanna Amann (ICSI – Berkeley, US)

Joint work of Johanna Amann, Robin Sommer
Main reference J. Amann, R. Sommer, "Providing Dynamic Control to Passive Network Security Monitoring", in Proc. of the 18th Int'l Symp. on Research in Attacks, Intrusions, and Defenses (RAID'15), LNCS, Vol. 9404, pp. 133–152, Springer, 2015.
URL http://dx.doi.org/10.1007/978-3-319-26362-5_7

Passive network intrusion detection systems detect a wide range of attacks, yet by themselves lack the capability to actively respond to what they find. Some sites thus provide their IDS with a separate control channel back to the network, e.g., by interacting with SDN capable hardware. In the past, such setups tended to remain narrowly tailored to the site's specifics with little opportunity for reuse elsewhere, as different networks deploy a wide array of hard- and software and differ in their network topologies. To overcome the shortcomings of such ad-hoc approaches we present a network control framework that provides passive network monitoring systems with a flexible, unified interface for active response, hiding the complexity of heterogeneous network equipment behind a simple task-oriented API. We give our experiences deploying our framework in a production network. Furthermore, we sketch future research directions that offload expensive low-level operations from software into network hardware.

4.2 Improving Network Security by SDN – OrchSec and AutoSec Architectures

Kpatcha Mazabalo Bayarou (Fraunhofer SIT – Darmstadt, DE) and Rahamatullah Khondoker

Main reference A. Zaalouk, R. Khondoker, R. Marx, K. M. Bayarou, "OrchSec: An orchestrator-based architecture for enhancing network-security using Network Monitoring and SDN Control functions", in Proc. of the 2014 IEEE Network Operations and Management Symposium (NOMS'14), pp. 1–9, IEEE, 2014.
URL http://dx.doi.org/10.1109/NOMS.2014.6838409
Main reference R. Khondoker, P. Larbig, D. Senf, K. Bayarou, N. Gruschka, "AutoSecSDNDemo: Demonstration of Automated End-to-End Security in Software-Defined Networks", in Proc. of the 2nd IEEE Conf. on Network Softwarization (NetSoft'16), pp. 347–348, IEEE, 2016.
URL http://dx.doi.org/10.1109/NETSOFT.2016.7502404

According to statistics of Deutsche Telekom [1], the number of network attacks per month has increased from 100,000 to 550,000 within 12 months (June 2015 – June 2016). Traditional defense mechanisms that are based on the strategy to automatically detect and manually mitigate attacks are deemed inefficient especially in the context of Industrie 4.0 applications. The concept of Software-Defined Networking (SDN) is based on the separation of the control plane from the data plane of network entities, whereas an SDN controller (representing the control plane) takes decisions based on forwarding rules, routers, switches, etc. (representing the data plane) forward the data accordingly. The planes communicate with each other by an open interface, such as OpenFlow, so that the data plane can directly be programmed. Among others, these centralized monitoring and control features of SDN can be adopted to detect and mitigate network attacks automatically. Towards this, two architectures named OrchSec

[2, 3] and AutoSec [4], have been developed by Fraunhofer SIT. While OrchSec detects and mitigates network attacks, such as DDoS, automatically in a reactive manner, AutoSec takes proactive actions, such as dynamically configuring both the clients connected to a network and the devices forwarding the data, to prevent the networks from being attacked successfully. OrchSec and AutoSec have been integrated and tested in SDN-enabled/SDN-only hardware devices from major switch vendors, such as Huawei, HP, and Cisco.

References

1 DTAG. *Overview of current cyber attacks on Deutsche Telekom AG (DTAG) sensors*. http://www.sicherheitstacho.eu/?lang=en, accessed on 04.08.2016

2 Adel Zaalouk, Rahamatullah Khondoker, Ronald Marx, Kpatcha M. Bayarou. *OrchSec: An orchestrator-based architecture for enhancing network-security using Network Monitoring and SDN Control functions*. NOMS 2014:1–9

3 Adel Zaalouk, Rahamatullah Khondoker, Ronald Marx, Kpatcha M. Bayarou. *OrchSec Demo: Demonstrating the Capability of an Orchestrator-based Architecture for Network Security*. Academic Demo, Open Networking Summit 2014 (ONS 2014), Santa Clara, USA, 3-5 March 2014

4 Rahamatullah Khondoker, Pedro Larbig, Daniel Senf, Kpatcha Bayarou, Nils Gruschka. *AutoSecSDNSemo: Demonstration of Automated End-to-End Security in Software-Defined Networks*. IEEE NetSoft 2016, 6-10 June 2016, Seoul, South Korea

4.3 SDN: A Network Economics Inflection Point

L. Jean Camp (Indiana University – Bloomington, US)

Joint work of Kevin Benton, L. Jean Camp, Martin Swany

Main reference K. Benton, L. J. Camp, "Firewalling Scenic Routes: Preventing Data Exfiltration via Political and Geographic Routing Policies", in Proc. of the 2016 ACM Workshop on Automated Decision Making for Active Cyber Defense, pp. 31–36, ACM, 2016; pre-print available from author's webpage.

URL http://dx.doi.org/10.1145/2994475.2994477

URL http://www.ljean.com/files/SAFECONFIG2016.pdf

Main reference C. Hall, D. Yu, Z.-L. Zhang, J. Stout, A. M. Odlyzko, A. W. Moore, L. J. Camp, K. Benton, R. J. Anderson, "Collaborating with the Enemy on Network Management", in Proc. of the 22nd Int'l Workshop on Security Protocols, LNCS, Vol. 8809, pp. 163–171, Springer, 2014; pre-print available from author's webpage.

URL http://dx.doi.org/10.1007/978-3-319-12400-1_15

URL https://www.cl.cam.ac.uk/~rja14/Papers/spw14-08-Anderson.pdf

BGP enables as a network of networks, and is also a network of trust. The most clear instantiation of that trust is the updating of router tables based on unsubstantiated announcements. The positive result of this trust is that the network can be extremely responsive to failures, and recover quickly. Yet the very trust that enables resilience creates risks from behavior lacking either technical competence or benevolence. Threats to the control plane have included political interference, misguided network configurations, and other mischief. BGPSEC has been proposed to resolve this, but the economics of path validation are the opposite of incentive aligned.

SDN offers an new approach to economics of networking. To show that this inflection point can improve network-wide security, we constructed a proof-of-concept. This proof of concept translates a series of route updates into a RIB, which is then converted to a flow information base (FLIB). The FLIB then can be subject to arbitrary analysis to defeat different types of attacks. For example, content-leaking misdirection attacks via incorrect

routing announcements could become immediately identifiable and individual networks could defend themselves from remote actors.

4.4 Network Security Management for Trustworthy Networked Services

Georg Carle (TU München, DE)

Joint work of Georg Carle, Cornelius Diekmann, Paul Emmerich, Sebastian Gallenmüller, Oliver Gasser, Nadine Herold, Matthias Wachs

When looking back to the previous research area of active and programmable networks 20 years ago, today's architecture of SDN-based networks can be seen as an evolution of these approaches. Our network security management approach combines different methods and components: Tools for automated and reproducible experiments allow automated load and penetration tests of real software and automated mitigation [1], [2]. Internet-wide measurements [3] provide a range of data that can be used in the testbed. Formally verified tools that allow to generate SDN flow tables and firewall rules from high-level specifications [5], and also allow to translate configurations of legacy devices into the same high-level specifications [4].

References

1 Emmerich, Paul and Gallenmüller, Sebastian and Raumer, Daniel and Wohlfart, Florian and Carle, Georg. *MoonGen: A Scriptable High-Speed Packet Generator*. ACM SIGCOMM Internet Measurement Conference (IMC) 2015, Tokyo, Japan, October 2015
2 Wachs, Matthias and Herold, Nadine and Posselt, Stephan-A. and Dold, Florian and Carle, Georg. *GPLMT: A Lightweight Experimentation and Testbed Management Framework*. Passive and Active Measurement: 17th International Conference, PAM 2016, Heraklion, Greece, March 2016
3 Wachs, Matthias and Herold, Nadine and Posselt, Stephan-A. and Dold, Florian and Carle, Georg. *Scanning the IPv6 Internet: Towards a Comprehensive Hitlist*. 8th Int. Workshop on Traffic Monitoring and Analysis TMA 2016, Louvain-la-Neuve, Belgium, April 2016
4 Diekmann, Cornelius and Michaelis, Julius and Haslbeck, Maximilian and Carle, Georg. *Verified iptables Firewall Analysis*. IFIP Networking 2016, Vienna, Austria, May 2016
5 Diekmann, Cornelius and Korsten, Andreas and Carle, Georg. *Demonstrating topoS: Theorem-Prover-Based Synthesis of Secure Network Configurations*. 2nd International Workshop on Management of SDN and NFV Systems, manSDN/NFV 2015, Barcelona, Spain, November 2015

4.5 RADIator – An Approach for Secure and Controllable Wireless Networks

Radoslaw Cwalinski (BTU Cottbus, DE)

Joint work of Radoslaw Cwalinski, Hartmut König
Main reference R. Cwalinski, H. König, "RADIator – An approach for controllable wireless networks", in Proc. of the 2nd IEEE Conf. on Network Softwarization (NetSoft'16), pp. 260–268, IEEE, 2016.
URL http://dx.doi.org/10.1109/NETSOFT.2016.7502421

Wireless local area networks (WLANs) became an essential part of todays enterprise network infrastructures. Due to the use of a shared medium – the electromagnetic waves – for transmitting data, wireless networks are inherently exposed to diverse attacks, such as for example Denial of Service (DoS) attacks at different network layers.

In the talk, we propose a software-defined networking architecture for enterprise wireless local area networks. In our architecture, the access point's (AP) management tasks, including beaconing, client authentication and association, are performed by the central controller instead of by the distributed wireless APs as in traditional networks. The goal is to provide a framework that exposes tools and methods for centralized, fine-grained inspection and processing of 802.11 frames and enable network applications to run in the central controller.

We present our architecture together with examples of controller-based applications that we are currently working on. These applications, such as centralized traffic inspection, anomaly detection, WLAN topology and interference recognition, wireless client geolocalization and client fingerprinting help to optimize and secure the WLAN. We introduce a "trust level"-based access control for wireless clients that uses geolocation information ("where you are"), device fingerprinting ("what you have"), anomaly detection ("what you do") and user credentials ("what you know") to take access decisions, set routing rules or trigger alerts.

4.6 The THD-Sec network security experimental testbed

Hervé Debar (Télécom & Management SudParis – Evry, FR)

The THD-Sec platform is an experimental environment dedicated to network security. It aims at enabling multiple attack and defense scenarios to provide experimental validation of new ideas for network defense. It includes classic IT technologies and interfaces to SCADA protocols. Examples of use of the platform have been published in [1] and [2].

References

1 Sahay, Rishikesh and Blanc, Gregory and Zhang, Zonghua and Debar, Hervé. *Towards Autonomic DDoS Mitigation using Software Defined Networking*. SENT 2015, Feb 2015, San Diego, Ca, United States. Internet society
2 Fabre, Pierre-Edouard and Debar, Hervé and Viinikka, Jouni and Blanc, Gregory. *ML: DDoS Damage Control with MPLS*. NordSec 2016:101–116

4.7 Security in ICS Networks

Tobias Limmer (Siemens AG – München, DE)

Many Industrial Control System solutions have a similar networking topology for which a common deployment practice has developed. As security standards increasingly gain attention, those deployments need to be adapted to new security requirements. This does not only apply to the design of the solution, but also to documentation, implementation, and verification practice. This talk presents an overview of the common deployment practice, security requirements, and open questions.

4.8 Authentication and Authorization in Wired OpenFlow-Based Networks Using 802.1X

Michael Menth (Universität Tübingen, DE)

Joint work of Frederik Hauser, Michael Menth, Mark Schmidt

802.1X is the most widely used authentication and authorization protocol in wired LANs. However, in OpenFlow-based networks, mainly MAC-address-to-identity mapping and web frontend based mechanisms are used which are highly insecure or cumbersome and little flexible, respectively. We propose to integrate the 802.1x authenticator in a network application such that it can support also others than RADIUS-based authentication resources. Further, a network-wide session database is maintained which enables identity-based network control. The authenticator is a network function that can be virtualized and well scaled. Most importantly, the approach is compatible with current infrastructures such as network clients and existing RADIUS-based authentication resources.

4.9 Robust Policy Checking

Christian Röpke (Ruhr-Universität Bochum, DE) and Thomas Lukaseder (Universität Ulm, DE)

The complexity and strategic position of SDN controllers in the network make them a rewarding target for attacks. Taking over an SDN controller means complete control over the network infrastructure. Despite their importance and their value, both for network operators and attackers alike SDN controllers are not secured properly against attacks in their current state. The complex structure of SDN controllers that also offer the possibility of including third party applications makes them hard to secure. Policy checkers are able to verify the compliance of the network set-up against a set of policies and can therefore serve as a warning system whether a controller is compromised. However, current policy checkers are usually placed close to the SDN controller on the same machine. Prior research shows that identifying a compromised SDN controller as such can therefore be circumvented by an

attacker. We discuss our ideas on different possible ways to integrate policy checkers in the network independently of SDN controllers. This makes policy checking more robust against a compromised control plane.

4.10 Initial Measurements on Delay Issues within SDN WAN-Scenarios

Thomas Scheffler (Beuth Hochschule für Technik – Berlin, DE)

Joint work of Thomas Scheffler, Awono Ngono, Gabrielle Nelly

Current SDN deployment focuses on data-centers where large content-providers have shown the value of the technology. As the technology matures and equipment becomes more readily available, other deployment areas may become interesting. Our work focuses on the use of SDN technology in Wide Area Networks (WANs). It has been shown before by others [1] that a small number of controllers could serve a large geographic area, such as the Internet2. SDN-WAN deployments would naturally contain certain controller-switch paths that facilitate high propagation delay.

Assuming that such networks use reactive flow instantiation, the following condition holds: whenever traffic reaches the switch, for which no match could be found in the flow table, there exists the need to forward OFP 'packet-in' packets to the controller. These OFP packets will have to be send over a high-delay link and may have a tendency to queue up, if several such events occur in rapid succession. We expect that a high switch-controller delay may alter the behaviour of the network and may have consequences to the end-to-end connections represented by these flows.

In the talk we present our testbed that allows us to introduce a variable, controlled delay between the SDN switch and controller. Our experiments show that in certain circumstances a high switch-controller delay leads to a large number of OFP packets forwarded to the controller. Current SDN switches simply forward all incoming packets for an unknown flow to the controller. One or several high-bandwidth flows thus flood the switch-controller link with many unnecessary OFP packets that still need to be forwarded to and processed by the controller. Since these packets are forwarded via a high-delay link, a large number of packets are already in flight, before a control message can reach the switch. This could potentially lead to an increased work-load on the controller, saturation of the switch-controller link, increased packet-forwarding delay, and the introduction of novel Denial-of-Service scenarios. We also found that delay values higher than 150ms affect TCP connections, represented by the flows, causing additional retransmission of packets to reach the network.

References

1 Brandon Heller, Rob Sherwood and Nick McKeown. *The Controller Placement Problem*. Proceedings of the First Workshop on Hot Topics in Software Defined Network (HotSDN'12), Helsinki, Finland, 2012

4.11 Party's Over – Why we are not only late to the SDN party

Alexander von Gernler (genua GmbH – Kirchheim bei München, DE)

Discussions about SDN are nice, but what if our insights will later on not be needed by the real world, because they have found better alternatives or doing it on their own no matter what we recommend? In this talk, I analyse the needs of several potential SDN users, namely data centers, company networks, and university networks. Data centers will mostly undergo a market consolidation, leaving out barely more players other than the cloud services of the Big Five companies, among them Amazon AWS, Google, and Microsoft Azure. They most likely will not be in dire need of our insights generated at Dagstuhl, as they have enough manpower and resources to just do it on their own.

Company networks, on the other hand, will undergo a transformation getting much leaner, following ideas like Google's BeyondCorp. Thus, SDN will not be of great importance here as well. What is left are university networks. They are often open-minded and will adapt or at least try out new ideas conceived by science. But then again, they are a really small market, so the impact of our ideas will be limited if only used in a university context.

5 Working Groups: The Good and the Bad of SDN

5.1 What benefits more? Attack Surface or Opportunity for Defense?

Kpatcha Mazabalo Bayarou (Fraunhofer SIT – Darmstadt, DE)

Joint work of Kpatcha Mazabalo Bayarou, Georg Carle, Sven Dietrich, Paulo Jorge Esteves-Veríssimo, Mattijs Jonker, Thomas Lukaseder, Michael Meier, Christian Röpke, Thomas Scheffler, Han Xu

SDN definitely increases the attack surface and the standards notoriously lack security mechanisms, e.g., for authorization which are BAD. On the other hand, SDN provides means to implement new security features faster and introduce them into the system in cases that were not possible earlier which are GOOD. Detecting attacks may therefore become a lot easier and reliable.

So which of the two aspects is more relevant and how will the final balance be? The working group discusses the two aspects by considering what is bad or good for the attackers' perspectives. The same consideration is made with regard to the defenders' perspectives. For this discussion the members of the group come up with the consideration of the limitations that may face both sides depending on which aspect/case is under consideration i.e. the discussion on bad or on good.

The discussion on the BAD relates to the advantage that the attacker gets from the SDN technology. The centralized architecture of SDN, lack of defenders expertise, and immature technology could benefit the attackers. For example, the introduction of malicious controller apps may allow for wider impact of the attack.

The discussion on the GOOD relates to the advantage of SDN for defenders and the limitation SDN poses to attackers. The centralized architecture of SDN which brings global view of networks, open hardware interfaces, and central control might benefit the defenders.

For example, open hardware interface empowers developers and network operators to create tailored security solutions.

What is the final balance? Finding attack surfaces that the SDN brings, is the precondition to defend against them. When usable, affordable and standard solutions will be provided against the attack surfaces, then opportunities for defense will be increased as the defender will be able to create innovative protection mechanisms using SDN by shifting the focus from protecting the SDN itself.

5.2 Standardisation & Transparency

Radoslaw Cwalinski (BTU Cottbus, DE) and Hartmut König (BTU Cottbus, DE)

Joint work of Johanna Amann, L. Jean Camp, Georg Carle, Radoslaw Cwalinski, Marc C. Dacier, Jan Kohlrausch, Hartmut König, Thomas Scheffler

The goal of the working group was to discuss the benefits and disadvantages of standardization and transparency in Software-Defined Networks. On the one hand, with SDN/OF networks may converge to one standard and a few (open) implementations that are easier to secure or fix than the myriads of diverging solutions. On the other hand, monoculture is bad if successfully attacked.

Starting with the positive side of standardization the members of the working group identified the following aspects. First, standardization of protocols for controlling network devices mitigates the risks of erroneous configurations. Ideally, network devices operate with open interfaces, avoiding vendor lock-in and reducing costs. Standardization also brings more players into the game thus allows for competition whereas the current non-standardization create vendor lock-ins and software solutions that are not future-proof. Standardized interfaces allow network monitoring to use networking systems in an unprecedented way, i.e. to filter information that they do not need.

The group members recognized also the advantages of transparency which is particularly critical for routing and security applications. Transparency helps with testing, including penetration testing and fuzzing. It also allows conformance testing by different organizations with open test suites and open, public test results. The point is that although vendors claim to be standard compliant, it tends to be a false promise which cannot be easily verified without public test suites and public test results.

On the bad side, the group participants agreed that standardization is subject for manipulation for organizations with high resources. Complexity of standardization is a proven way to decrease the interoperability in practice thus increase opportunities for a vendor lock-in. Additionally, complex standard interfaces are hard to set up and to manage. They also can come with "standard vulnerabilities". These vulnerabilities might therefore affect an even larger number of standardized systems. Network monoculture of such standardized systems may make it easier for attackers to compromise the system's security. Further, current standards are often not suited for SDN, e.g., the standards of PKI for SDN are inappropriate. They can offer a false feeling of authentication and an illusion of security.

SDN will always need to interact with the legacy world. This interaction sets limitations to the security benefits of SDN. The challenges of BGP will not disappear with SDN – important threats like BGP prefix hijacking remain difficult to deal with. In addition, the presence of legacy middleboxes can also break many SDN-based security mechanisms. Debugging

methods from legacy networks may be affected by SDN too e.g., ping may not follow the same path as http. Generally speaking, SDN programming may be influencing traffic in a complex way. The conclusion is that SDN promises network transparency but also challenges it.

The participants agreed that standards are often battles for finite resources. Increasingly, the standards become more complex and burden developers which leads to increased complexity at the software level. The sad truth is: security is traditionally sacrificed for interoperability.

Finally, the separation of organizations served on an airport has been presented as an use case to demonstrate the benefits of SDN. Today the separation is mostly done with MPLS which is limited and cumbersome to configure. Using SDN the isolation can be done in a convincing and straightforward way. Another example presented was the isolation of flows within an aircraft and between an aircraft and ground data centers involving different organizations: aircraft manufacturer, engine manufacturer, airline, maintenance organization, airport.

5.3 Flexibility and Adaptability for Attackers and Defenders

Boris Koldehofe (TU Darmstadt, DE)

Joint work of Boris Koldehofe, Simin Nadjm-Tehrani, Rene Rietz, Robin Sommer, Jens Tölle, Emmanuele Zambon

Preface: Some of the given statements are not exclusively valid for SDN. The advantages and disadvantages can occur with other advanced network management technologies as well. Standardized and widely-used approaches will intensify opportunities and risks.

Starting with the potentially problematic aspects of SDN usage the members of the working group identified the following challenges, most of them a cause of increased complexity:

- Code for managing and configuring SDN capable switches may come from various sources, and some of them may contain malicious contents.
- Networking devices may have technical capabilities which are not used by most of the users. So it is not transparent to hosts what the actual network configuration is.
- If more than one user is allowed to configure the system, even with good intentions there will be unknown side effects taking the system to places the service provider did not imagine.
- The flexible updates creates a need for much more complex access control systems that are hard to manage, and add to the complexity of the overall system.
- The notion of normality is harder to define in an SDN that is programmable simply due to larger degrees of freedom, and hence detection of abnormal events gets harder. Attackers can use this "confusion" of conception to hide the attack steps. The need for flexibility will mandate for more extensive interpretation of network data (i.e., looking at/parsing the application layer). This will increase the attack surface in both SDN switches and controllers.
- Attackers may get the same capabilities as the operators once they breach the trust management system – and they will exploit it.

All in all, attackers can actually control the operations in arbitrary ways, they can confuse or blind the defenders, or create inconsistencies. They are able to gather a global view of the network (and a more fine-grained too) from a single location. They will be able to exploit the additional complexity brought in by the flexibility (e.g., code exploitation on switch-side and controller-side).

The flexibility makes it harder for the defender of SDNs. Because of dynamic configurations, it is more difficult for a human to tell if the current/past configuration is intended/correct. The more user-friendly tools get, the less humans are able to do the job themselves and have a deep understanding of the underlying technology and protocols. The flexibility makes it hard to define meaningful policies for SDNs, e.g. which flows are affected by a specific network application and modified in a specific way. The flexibility provided by SDNs may exacerbate the conflicts between the objectives of networking teams vs. security monitoring teams.

The working group discussed also the positive aspects of application of SDN technologies. From the point of view of defenders, it gets easier to:

- do static and dynamic network isolation
- do fine granular authentication/authorization of clients
- enable active response (blocking, restricting), including deep inside the local network
- gain network overview, creating awareness on current security situation
- do adaptive monitoring (e.g., tell the switch that we don't want to see this particular flow (file transfer) anymore)
- do efficient network monitoring using in-network processing
- creating resilience: enable rate limiting or rerouting of traffic when under attack.

From an attackers point of view, the following attack-related activities get harder:

- Network reconnaissance
- Analysis of a properly separated network environments
- Man-in-the-middle attacks using spoofing (if there is a proper SDN concept, e.g., address configuration/resolution using SDN services)
- Takedown of a complete system (e.g., by limiting the attack to certain services)

Conclusion: All in all, what we see as the real added value of SDN to security is the ability to interact with switches and routers by means of APIs. These APIs can be leveraged for a number of security-related tasks, independently from the complete adoption of the SDN paradigm.

5.4 Too novel to be applied or the way out of security ossification?

Tobias Limmer (Siemens AG – München, DE)

Joint work of Marc Eisenbarth, Felix Erlacher, Frank Kargl, Thomas Kemmerich, Tobias Limmer, Ramin Sadre, Sebastian Schmerl, Bettina Schnor, Radu State, Alexander von Gernler

SDN is a novel technology and may solve several problems that are surfacing in current network topologies. Increasing heterogeneity, caused by new initiatives such as Bring Your Own Device (BYOD) or developments in the area of Internet of Things (IoT), or highly dynamic network changes required by virtualization are just a few examples.

To control the effects of those new developments, more fine-grained control is necessary as is currently supported by legacy networking equipment. For example, the augmentation of traditional firewalls that allows them to examine and filter intra-subnet traffic may help to protect potentially untrusted endpoints from each other. SDN supports this use case by introducing a common transparent interface to networking devices for network security mechanisms. Using this standard interface, software and devices from different vendors may become interoperable and may be managed within one environment. However, the current state of available standards, such as OpenFlow, is not promising here. It can be easily seen that those standards and related regulations are still immature, as important parts are not defined yet. In the case of OpenFlow, northbound interfaces are not standardized yet, and available network apps typically disregard security completely.

The new possibilities in the security area are based on the flexible architecture of SDNs. This fact results in configurations and network topologies that may become very complex. From a technical point of view, a diverse set of problems arises here: SDNs usually should distribute components within the network to ensure reliability. What happens if multiple controllers issue conflicting instructions to network devices? In what way should controllers prevent problematic situations caused by multiple interacting networking apps that have been downloaded from a central app store? What happens if a network is segmented in multiple parts, and newly appearing devices need to be boot-strapped to be integrated into the network? Many security applications within SDNs also rely on packet forwarding to centralized components which may analyze those packets. On the one hand, SDNs are supposed to make a network more efficient, but on the other hand, new features may lead to uncontrolled network link congestions which may require even higher data rates compared to traditional networks. The complexity of SDNs may also impact compliance certifications in the banking sector or safety regulations in the area of Operational Technology (OT). These questions are still largely unresolved and need to be addressed before SDNs are deployed in this flexible operation mode.

Still, many large Internet companies and ISPs show much interest in SDN deployments, and several of those make already use of SDNs. In the current state, much expert know-how and many customizations are necessary to successfully deploy SDN and benefit from its features. Facebook, as an example, already has an SDN-based deployment method for big data centers. ISPs may benefit from a common framework of all network devices which supports a common language to express network policies and rules. This would allow providers to simplify policy compliance and configuration, and may even open new business opportunities such as customers who could upload apps to their provider's infrastructure for customized network features, such as DDoS protection, QoS, or packet filtering. Due to open standards and one common environment, those network apps can be sandboxed by the underlying controller, allowing to separate network logic and security.

Instead, we may also continue to rely on proven and well-established security technologies like firewalls or intrusion detection systems that we know how to handle. If network topology and devices are chosen carefully, most of the features that can be realized with SDNs are also available within traditional networking environments. Furthermore, SDNs will only be able to fully automatically control and manage the simplest networks – customization and management by network experts will still be necessary in many cases. But what about networks that constantly face changing requirements from the business side, technical problems caused by evolved network topologies with devices from different vendors in different versions? Here, SDN may provide a solution due to its capabilities to standardize interfaces and features across vendors and network devices.

5.5 Is SDN more complex or simpler?

Claas Lorenz (genua GmbH – Kirchheim bei München, DE)

Joint work of Dieter Gollmann, Peter Herrmann, Hartmut König, Claas Lorenz, Michael Menth, Björn Scheuermann

The concept of SDN promises a reduction in complexity by splitting networks into a dedicated data plane and a logically centralized control plane. When explaining concepts like routing, the software approach in SDN seems much more simple than the distributed algorithms and protocols in classical networks, since it can just be represented as a simple graph problem. This narrative is stressed by two aspects that are hidden in the simplicistic model of SDN regarding the controller as a single entity rather than a distributed system. The need for scalability and operational requirements, e.g., concerning fault tolerance, enforce a distributed approach. Additionally, the realization of the control plane completely in software raises issues about its algorithmic complicateness. This is due to the additional requirements that were not imposed on classical networks, but are now thinkable in SDN. While this is a unique selling point in terms of possible features, it raises serious concerns for security, as it opposes simplicity which is a key design principle for building secure systems.

State-of-the-Art controller implementations suffer a tremendous feature bloat which is most likely buggy and rather untested. The same problem occurs with switches, which are often legacy equipment,are enriched with an OpenFlow interface. The simplicity, as intended by the SDN paradigm, is not very common in practice which might be a result of the consortial standardization model leading to hard fights between financially potent parties and feature rich compromises in standards and implementations. Nevertheless, there exist industry-grade whitebox switches as well as simple, lightweight controller implementations. For the price of providing less features, the realization of SDN using simple and possibly less attackable components is possible. Nevertheless, if an advanced feature set is required the controller must be designed as a distributed network operating system with security enforcement mechanisms in place analogous to traditional operating systems. An example trait would be the distinction between a kernel and a user land with well-defined interfaces and access control.

Flows as data model in a switched network are much simpler notions than layered packets in traditional routed networks. This may help to define a general structural core while providing powerful functionality. If this core could then be standardized and implemented very narrowly it is likely to be well designed, broadly tested, and hardened properly. On the other hand, the separation of data and control plane creates different views and, with emphasis on their consistency, makes the creation of a wholistic security solution a tough challenge. Even though, this distinction makes the decomposition of components easier and therefore better testable. Also, security patches for the control plane become more feasible.

Besides the defense of the SDN itself, it can be used to simplify mitigation of attacks that are commonly seen in classical networks. Attacks like ARP flooding or DHCP spoofing can be tackled in a simple and effective manner with SDN. In addition, every switch may provide firewalling functionality helping to achieve a defense in depth.

All in all, SDN introduces numerous challenges regarding complexity and simplicity of the system. It has the potential to be simple but making it simple is quite complex. The decomposition of components is easy, but their secure reassembly remains challenging. Therefore, a self-limitation regarding the necessity of features must be taken into consideration to allow a simple and secure design, implementation, and operation of a Software-Defined Network.

5.6 The Good and the Bad of Centralization in SDN

Christian Rossow (Universität des Saarlandes, DE)

Joint work of José Jair C. de Santanna, Issa Khalil, Evangelos Markatos, Michael Menth, Christian Rossow

By design, SDN centralizes many networking aspects that traditionally might have been decentralized. For example, SDN-driven networks may offer a centralized location to access or steer the data, control, and management plane. Furthermore, SDN-driven networking algorithms can assume a centralized data model, which was not possible in traditional networking. Since this is a radical change in the way we think of networks, we have investigated in our working group pros and cons implied by the centralization aspects of SDN.

First, a centralized architecture and network management creates a single point of failure which downgrades the resilience given by a distributed system. It is debatable whether traditional networks do not already offer single point of failures, but SDN adds some additional centralization points that might be exploited by an (i) internal attacker that suddenly has a central place to monitor and manipulate the network or (ii) by an external adversary that compromises vulnerable SDN components. This requires further thoughts on how SDN can be protected against such attacks.

Second, it may happen that the centralized decision engine of SDN adds a new type of denial-of-service (DoS) vector. For example, an attacker might be able to overload the controller with unknown flows that require constant decision makings. On the other hand, the centralization of SDN allows to more effectively tackle existing types of DoS attacks, as it has a global view of the network topology and can correlate this information with the traffic analysis for more reliable attack detection results. The two areas bear interesting research questions that should be investigated further.

Third, an important aspect is how SDN signaling is organized, in-band or out-of-band. If both the data and the control plane share the same (physical or logical) network segment (in-band signaling), the control plane may also become corrupted if the data plane breaks. As a consequence, out-of-band signaling schemes should be explored further to allow an easier recovery.

Fourth, scalability is a key feature of centralized systems. SDN involves a few critical parts that may become bottlenecks, however. For example, the flow tables may fill, so that the hierarchy of the networks requires careful thinking. In addition, if a layer of redundancy or load balancing is added (e.g., in terms of multiple controllers), suddenly there is the need for communication to avoid any possible state or decision inconsistencies. These aspects motivate further research how the centralized parts should be designed in a scalable fashion.

Fifth, although SDN increases the network complexity and the plentitude of intertwining algorithms may emit possibly contradicting policies, we are convinced that it is especially the centralization that plays in our hands to resolve such inconsistencies. Reacting to and removing such policy inconsistencies is much easier in a centralized network, such as SDN. This has positive implications on many types of policies, such as centralized routing algorithms, firewalls, or network monitoring methodologies.

To sum up, the centralization imposed by SDN indeed creates new challenges, but the benefits are clearly predominant. However, it is important to address the open research questions in this regard to ensure security and resiliency of the centralized SDN aspects.

6 Working Groups: Research Directions

6.1 Research Directions: Methods, Policy, and Attacker Model – Assessing and Improving the Security of SDN Networks

Georg Carle (TU München, DE)

Joint work of Georg Carle, Sven Dietrich, Dieter Gollmann, Bettina Schnoor, Peter Herrmann, Christian Röpke

When assessing suitable approaches for specifying security goals for SDN, it was identified that existing methods include natural language approaches, such as the ones used in ISO 27000, Common Criteria, BSI Base Protection Catalogue, and also formal approaches, as part of Linux iptables, Unified Modeling Language (UML), Security Policy Languages, and BAN logic from the protocol analysis field. It was identified that an important goal is to automatically derive secure SDN configurations. That requires extensions of the state-of-the-art methods, by providing additional information elements for the full range of components of SDNs, representing all states of SDN network elements. There is also a need for new tools that are capable with dealing with this additional information.

Methods to assess the security of SDN networks range from penetration testing to formal analysis, such as using policy checkers. Penetration testing has several limitations, such as the limited coverage of the system. It may also be difficult to identify the problems that tests do not find. In particular, the outcome of various tests may depend on the state a specific SDN component is in, which may depend on past input via different network interfaces. Policy checkers allow one to identify a case where a set of policy rules violates a set of security policies. However, if the policy set is incomplete, it is possible that certain violations would not detected. On the other hand, with penetration testing such violations that are not detected by formal methods may indeed be detected.

When assessing what current policy checkers cannot detect in SDN networks, it was identified that concurrency violations are an important problem in SDN, as this may lead to policy or invariant violations, such as blackholes, forwarding loops, or non-deterministic forwarding [1].

Methods to provide a trust base for SDN include providing a security kernel inside the SDN controller [2], which are able to distinguish between various types of SDN controller applications. For example, in the case of coexistence of a firewall and a load balancer application on the controller, the firewall application would have priority over the load balancer application.

Concerning relevant attacker models, it was identified that related work, such as [3], provides a highly useful taxonomy of attacker models. In order to prevent that possible attacks may be successful, one can consider an approach in which the different states a network may be distinguished. That means identifying good states in which the known attack cannot be successful while avoiding bad states. For the latter bad states, it is known that attacks can be successful.

Overall, it was identified that the differences of SDN to conventional networks make it very hard to ensure the security of SDNs. This is a consequence of the additional complexity of SDN, in which controllers change the configuration of the switches, allowing for a variety of automated reconfigurations. This makes attacks possible in which an attacker causes a reconfiguration to occur that leads to the desired outcome. For example, an attacker may create legitimate but dubious traffic, thereby causing the controller to regularly reconfigure the switches.

All approaches that allow one to handle the increased complexity are considered to be highly useful. They ensure that certain SDN applications can only influence certain flows. By applying the concept of network isolation, SDN enables network slicing and Virtual Network Operators.

References

1 Ahmed El-Hassany, Jeremie Miserez, Pavol Bielik, Laurent Vanbever, and Martin Vechev. *SDNRacer: Concurrency Analysis for Software-Defined Networks*.PLDI'16, Santa Barbara, CA, USA June 2016

2 Phillip Porras, Seungwon Shin, Vinod Yegneswaran, Martin Fong, Mabry Tyson, Guofei Gu. *A security enforcement kernel for OpenFlow networks*. First workshop on Hot topics in software defined networks HotSDN 2012, Helsinki, Finland, August 2012

3 Diego Kreutz, Fernando M. V. Ramos, Paulo Veríssimo. *Towards secure and dependable software-defined networks*. Workshop on Hot Topics in Software Defined Networking HotSDN 2013, Hong Kong, August 2013

6.2 Research Directions: Secure Operations in SDN-based Environments

Marc C. Dacier (QCRI – Doha, QA)

Joint work of L. Jean Camp, Marc C. Dacier, Jan Kohlrausch, Tobias Limmer, Michael Meier, Simin Nadjm-Therani, Ramin Sadre, Thomas Scheffler, Sebastian Schmerl, Radu State

On Thursday, September 8, 2016, one of the themes debated by the participants in a parallel session was oriented towards the issues on how to securely operate an SDN-based environment. It led to a very lively discussion for several hours, the gist of it is summarised here below.

Before thinking of operating an SDN environment, a key question discussed by the team was related to the rolling out of SDN in an existing environment. There was a consensus to say that it was unlikely that (i) SDN would completely replace an existing, non SDN based, environment and that (ii) any deployment would have to take place in an incremental way. In both situations, namely transient phase of deployment and ongoing operation of a mixed environment (SDN and non SDN), it was felt that specific security concerns would have to be addressed since the promises of an homogeneous, well defined, centrally controlled SDN environment would not be present. There was the feeling within the group that such operational concerns were not properly addressed by existing solutions yet and that it would deserve some further research to lead to practical solutions.

The group generally agreed that SDN would not replace but instead complement the networking toolbox at the disposal of operators. Two specific use cases were discussed where SDN was seen as a, possibly, useful paradigm to use. The first one was related to the emerging "Bring Your Own Device" paradigm (BYOD) in which potentially compromised devices were dynamically added to the networking infrastructure. The need for a simple and clear mechanisms to enforce well defined policies for such devices was an argument in favour of an SDN environment. Indeed, if well done, SDN could be used to automatically implement concepts such as the quarantine of misbehaving devices, degraded – or fail safe – modes for the network in case of worm propagations, adaptive scrutiny of network flows to look for data exfiltration, etc...

The second use case discussed by the group was related to critical infrastructures or, more generally, so called "Operational Technology" (OT) environment, as opposed to "Information Technology" (IT). It was noted that, nowadays, whereas the OT department was in charge of running the OT infrastructure, its security still usually felt under the responsibility of the IT department. It was observed that in such deployment, SDN could help the IT department in improving the limited visibility they currently have and would make it easier for them to enforce, at the networking level, the needed security policies. A contrario, it was also acknowledged that OT environments are quite resistant to changes and a convincing argument had to be brought forward to implement such radical change which would, quite likely, require to replace most, if not all, routers and switches in these environments.

More generally, it was felt that, whereas SDN clearly has some claimed benefits, there was a need for a thorough economic study of the pros and cons which would take into consideration the possible negative effects on security and the supplemental costs associated with a reinforcement of the needed security tools.

The human dimension of the SDN impact on security was also discussed. Not only in the way its deployment could be a bridge between the IT and OT worlds, as discussed before, but also with the increased risks created by giving a lot of powers to the few (or sole?) administrators of the SDN controller. As we see more attacks due to insider, it was agreed that the risk of having a malicious administrator was not to be neglected and to be dealt with but more research was required to come up with a satisfactory solution. Along the same line, there was some fear expressed that the possibility of having various kinds of applications running in the controller to serve different purposes could lead to some serious organisational disputes if not properly anticipated. For instance, if two distinct departments (e.g. marketing and IT security) want each to have their own application in the controller, built on distinct requirements (e.g. quality of service vs. security), who would (i) detect possible inconsistencies between decisions made by these applications and (ii) decide which one to favour?

The problem of various applications, designed and developed by independent teams, running in the same controller is a very large problem that has been discussed at length by the team. It came out that there is a clear need for more research to be done in order to help the people running SDN platforms to decide not only if (i) a given application is secure in the first place (i.e. without any vulnerability, and not malicious) but, more importantly, if (ii) the addition of a new application to a controller where other applications are already running would not create security issues due to the composition of the decisions made by each application independently. Is it possible to prove, by construction, that, assuming each application is "secure", the software resulting from the composition of all these applications remains secure? This was seen as an important open research area.

Finally, it was expected that most of the problems that the domain of network operational security has been dealing with in the past would, could or should be revisited in the sense that the introduction of SDN was changing the attack surface that people had been used to consider when looking at distributed systems. For instance, the existence of a common controller used for two networks separated by a firewall could open the door for new techniques to circumvent the firewall (if SDN was not correctly configured). More generally, the presence of such common controller could be seen as a new way to implement well known covert channels. Also, an SDN environment, if not very securely configured, would offer lots of opportunities for new ways to launch denial of service attacks, to avoid detection by deep packets inspection devices etc.

All in all, it was felt that SDN could certainly help in improving the operational security

of a network environment but that many problems remain unsolved (i) to ensure that a given SDN environment would be secure by construction, (ii) to prevent malicious users (especially administrators) or applications from misusing such environment and (iii) to detect when such misuse would occur.

6.3 Research Directions: SDN-based Security

Frank Kargl (Universität Ulm, DE)

Joint work of Johanna Amann, Kpatcha Mazabalo Bayarou, José Jair C. de Santanna, Radoslaw Cwalinski, Marc Eisenbarth, Felix Erlacher, Marko Jahnke, Mattijs Jonker, Frank Kargl, Evangelos Markatos, Rene Rietz, Christian Rossow, Robin Sommer, Jens Tölle

The working group discussed how SDN would enable new forms of network security mechanisms to be envisioned, designed, and implemented, or how SDN would allow existing mechanisms to be implemented in a more flexible or interoperable way. For this, we first identified typical attacks where we assumed a potential for SDN-based mitigation mechanisms. Attacks we discussed included DDoS, reconnaissance, Man-In-the-Middle, malicious modifications of the network including any accidental misconfigurations, and malware-related attacks that we spread into initial infection, internal spread, Command & Control (C&C) communication and data exfiltration.

We then created a table where all these attacks were listed in relation to the common categorization of security mechanisms in prevention, detection, reaction, and forensics. For each of the resulting cells, we discussed how SDN would support or hinder the design of such security mechanisms.

The discussion results are depicted in Table 1. For the purpose of this text, we will only address what participants considered the most interesting ideas. In general, we identified that SDN enables mostly two types of capabilities that security mechanisms may make use of.

First, SDN and OpenFlow allow holistic control of network devices throughout all active network components. With this, mechanisms that inspect or filter traffic anywhere in the network become possible. Second, SDN offers a standardized interface for interacting with the network which would allow cross-platform security mechanisms that are not tight to a specific vendor.

For DDoS attacks, it should probably be investigated further how fine-grained filtering throughout the own network can help to either prevent such attacks or react to such attacks and filter out attack traffic and how this may be more effective than central filtering. However, this is probably mostly effective for egress filtering and therefore mitigating attacks that originate from your own network. Beyond, if we foresee the notion of "network apps", these may also be used to implement mitigation logic for a specific attack on your network. This mitigation logic could then be deployed in the network of your ISP in order to have a highly specific, fine-granular and customized filtering being created by the network operator executed within all ISP's devices.

Reconnaissance attacks may also be easier to detect with SDN. The assumption is that there is often no fixed central place to detect such attacks. Particularly if they stem from internal nodes, applying an IDS on your Internet gateway will not be effective and you would therefore deploy your IDS on many places inside your network. This may require substantial resources. We came up with the notion of Network Function Virtualization (NFV) of network security mechanisms like firewalls or IDSs/IPSs that would run on a central cloud server or

on cloud servers distributed in the network. You would then use the SDN functionalities to pre-filter traffic and forward the resulting streams or packets to the IDS for inspection. If there are suspicious activities being detected, you may even reduce filtering to inspect the traffic more intensively. Beyond, NFV of network security mechanisms in cloud datacenters would allow migrating the IDS or firewalls that monitor a certain critical virtual machine together with that virtual machines.

Regarding Man-In-the-Middle attacks, we discussed that SDN would allowing to quickly react to such attacks once they are detected. Hosts running such an attack could be quickly isolated and then investigated by forensic mechanisms. Regarding malicious or accidental modifications in the network, we think that SDN could help by having a central point where network configuration (including open flow tables) is accessible. Then, detecting inconsistencies and applying plausibility checks to this network state would allow detection of malicious modifications to routing, identification of unauthorized hosts, changes to network topology and many more such attacks.

At the same time, we also acknowledged that applying SDN in your network will, in general, make the configuration and the state of your network much more complex and thus detecting such attacks in the first place will become much harder. This is a general problem for network security in SDN-enabled networks: due to the high volatility and fine granularity of network configuration, it may be substantially harder to detect attacks. This applies also to other parts of this discussion, like the Man-In-the-Middle detection.

Regarding malware, we again identified a potential for applying NFV to have mechanisms like malware scanning or IDS being applied flexibly and scalable in the network. So if there is a malware outbreak and spread in one part of the network, resources can be allocated on your cloud servers to inspect particularly that traffic in that network segment. Likewise, if you have critical resources that get relocated to different parts of the network as part of cloud operations, the network security mechanisms may migrate together with them.

Next, SDN may also support easier containment of malware infections and spread. You may easily segment your network, e.g., triggered by the IDS or virus scanner having detected infections on some host. One idea was to even simulate the possible spread of a malware based on known SDN state. So if a malware is known to spread via a certain protocol, one could simulate which other hosts are reachable in a transitive way and then apply more stringent filtering and isolation to those hosts that are potentially infected.

Finally, malware may also be addressed by using SDN mechanisms to redirect the communication of infected machines with their C&C servers. This so called sinkholing would allow to redirect traffic to C&C's IP addresses to a security host where that traffic can be forensically analyzed, filtered, or even modified, e.g., to issue instructions to infected hosts. This will also allow to gather detailed statistics on infected machines.

Independent of those attacks, we also came up with the idea to use the isolation capabilities of SDN to create islands of personal devices within a network. Thus, all devices belonging to the same user – smartphones, tablets, smartwatches, laptops, etc. – would sit within the same island and could freely communicate with each other, including broad- and multicast discovery protocols, while external communication could be subject to a consistent security policy for that specific user. Overall, we considered SDN to be an interesting enabler for security mechanisms and could come up with a whole series of concrete ideas that we think would merit further investigations in future research projects.

Table 1 SDN-enabled security mechanisms.

	Prevent	Detect	React	Forensics
DDoS	Fine-grained filtering	Offloading certain filtering/detection operation at the switch level to be able to operate at line rate while extending inspection at more than netflow information	Using the whole network to react	Statistics, logging and packet inspection for better understanding how the DDoS works
Reconnaisence	(1) Using the whole network for filtering (2) hiding the network structure	Network Function Virtualization (NFV) for IDS, honeypot on-demand	NFV for IDS, honeypot on-demand (e.g., virtual deployment of a honeypot)	Statistics, logging and packet inspection
MITM (not at the application layer)	Fine-grained traffic control	(1) Detecting routing anomalies (may be harder in the presence of SDN, due to increased complexity) (2) Detecting forwarding correlations (also possible before SDN)	Quick isolation	(1) Negative: increased number of more complex states (2) Implement MITM for inspection
Misconfigurations & malicious modifications	Global policy with SDN	Consistency and plausibility checking on flow tables becomes more difficult due to increased complexity	Probing of network behaviour of dedicated resources (e.g., isolation of errors)	(1) Statistics, logging and packet inspection (2) Checking network invariants
Malware (initial) infection	NFV for virus scanner	(1) Using whole network for detection (2) NFV for IDS	IDS/quarantining potentially infected hosts	Logging, network-wide view to identify where the attack came from
Malware spread	(1) Pervasive possibility for isolation/segmentation (2) Segmentation may disrupt some services (e.g., NetBIOS)	(1) Using the whole network for detection (2) NFV for IDS	IDS/quarantining potentially infected hosts	Simulation of malware spread (feedbacks to better prevention and reaction)
Malware C&C	Sinkholing C&C	Netflow-like analysis	Sinkholing C&C	Redirecting C&C traffic for analysis
Malware data exfiltration		Detecting "NSA style" keyword exfiltration based on SDN logs	Modification/marking exfiltrated data	

6.4 Research Directions: Secure Architecture for SDN

Alexander von Gernler (genua GmbH – Kirchheim bei München, DE)

Joint work of Paulo Jorge Esteves-Veríssimo, Alexander von Gernler, Thomas Kemmerich, Issa Khalil, Boris Koldehofe, Hartmut König, Claas Lorenz, Björn Scheuermann, Han Xu

The working group dealt with the topic to find a secured architecture for SDN. Based on an exemplary diagram of an SDN setting we tried to identify security issues concerning single components, links, or functional elements of the SDN setting. We discussed whether there are applicable architectural patters and best practice experience. All participants agreed that there is a need for such architecture, but the time was too short to find a conclusive proposal. A solution of this problem requires deeper and long-term research. In our discussion, a number of questions have been raised which require further research activities. Among these were:

1. How to securely implement and deploy "network apps"? How to design the northbound interface so it is secure and expressive?
2. Complexity is an important issue in SDN. How can SDN solutions be simplified? How can SDNs scaled securely?
3. How to implement access control and authorization in SDN networks?
4. How can we protect the controller itself?
5. How can we secure the communication between controller & switches?
6. How can we perform intrusion detection and anomaly detection in SDNs?
7. How can we perform intrusion detection / resp. achieve SIEM functionality in the SDN context?
8. How differently do we have to deal with misbehaving/malicious clients?
9. How can we deal with misbehaving/rogue applications?
10. How to mitigate attacks?
11. What is the role of trusted hardware in switches? Is it needed for strong security?
12. How can you operate SDN in presence of untrusted HW components?
13. How do we ensure the software quality of the SDN infrastructure (controller, HW, ...)?

7 Final Plenary Discussion

7.1 Theses on SDN security

Hartmut König (BTU Cottbus, DE) and Radoslaw Cwalinski (BTU Cottbus, DE)

Joint work of all seminar participants

In the final plenary session of our seminar, the participants formulated the following theses regarding the security of SDN.

1. SDN is hard to define, one needs to be clear about assumptions and goals. SDN feature consolidation will come, but is not yet foreseeable.
2. The main advantage for SDN deployment will not be security. However, SDN creates a lot of security problems, many of which do not have a clear solution.

3. On the other hand, SDN enables new creative forms of security mechanisms – without being mandatory for them. Reaction possibilities to security incidents can be enhanced. One can use SDN for security even without full deployment of SDN in the network.
4. SDN security solutions demand a holistic approach including trusted computing base in network component. Secure software engineering will become more relevant for networks with SDN. Securing SDN, in particular network apps, requires substantial progress in software security and other fields, such as access control and policy definition.
5. Simple SDN solutions foster SDN security, but keeping SDN simple is complex!
6. Centralized controllers create many internal security challenges, e.g., "Packet INs" are considered harmful. More static uses of SDN are better for security.
7. There is no clear SDN/OpenFlow security roadmap.
8. Without security, SDN will not succeed!

Participants

- Johanna Amann
ICSI – Berkeley, US
- Kpatcha Mazabalo Bayarou
Fraunhofer SIT – Darmstadt, DE
- José Jair C. de Santanna
University of Twente, NL
- L. Jean Camp
Indiana University – Bloomington, US
- Georg Carle
TU München, DE
- Radoslaw Cwalinski
BTU Cottbus, DE
- Marc C. Dacier
QCRI – Doha, QA
- Hervé Debar
Télécom & Management SudParis – Evry, FR
- Sven Dietrich
City University of New York, US
- Falko Dressler
Universität Paderborn, DE
- Marc Eisenbarth
Arbor Networks – Waco, US
- Felix Erlacher
Universität Innsbruck, AT
- Paulo Jorge Esteves-Veríssimo
University of Luxembourg, LU
- Dieter Gollmann
TU Hamburg-Harburg, DE
- Peter Herrmann
NTNU – Trondheim, NO
- Marko Jahnke
CERT-BPOL – Swisttal, DE
- Mattijs Jonker
University of Twente, NL
- Frank Kargl
Universität Ulm, DE
- Thomas Kemmerich
Norwegian University of Science & Technology, NO
- Issa Khalil
QCRI – Doha, QA
- Hartmut König
BTU Cottbus, DE
- Jan Kohlrausch
DFN-CERT Services GmbH, DE
- Boris Koldehofe
TU Darmstadt, DE
- Tobias Limmer
Siemens AG – München, DE
- Claas Lorenz
genua GmbH – Kirchheim bei München, DE
- Thomas Lukaseder
Universität Ulm, DE
- Evangelos Markatos
FORTH – Heraklion, GR
- Michael Meier
Universität Bonn, DE
- Michael Menth
Universität Tübingen, DE
- Simin Nadjm-Tehrani
Linköping University, SE
- Rene Rietz
BTU Cottbus, DE
- Christian Röpke
Ruhr-Universität Bochum, DE
- Christian Rossow
Universität des Saarlandes, DE
- Ramin Sadre
University of Louvain, BE
- Thomas Scheffler
Beuth Hochschule für Technik – Berlin, DE
- Björn Scheuermann
HU Berlin, DE
- Sebastian Schmerl
Computacenter – Erfurt, DE
- Bettina Schnor
Universität Potsdam, DE
- Robin Sommer
ICSI – Berkeley, US
- Radu State
University of Luxembourg, LU
- Jens Tölle
Fraunhofer FKIE – Wachtberg, DE
- Alexander von Gernler
genua GmbH – Kirchheim bei München, DE
- Han Xu
Huawei Technologies – München, DE
- Emmanuele Zambon
SecurityMatters B.V., NL

Report from Dagstuhl Seminar 16362

Robustness in Cyber-Physical Systems

Edited by

Martin Fränzle[1], James Kapinski[2], and Pavithra Prabhakar[3]

1 **Universität Oldenburg, DE,** martin.fraenzle@informatik.uni-oldenburg.de
2 **Toyota Technical Center – Gardena, US,** jim.kapinski@toyota.com
3 **Kansas State University – Manhattan, US,** pprabhakar@ksu.edu

Abstract

Electronically controlled systems have become pervasive in modern society and are increasingly being used to control safety-critical applications, such as medical devices and transportation systems. At the same time, these systems are increasing in complexity at an alarming rate, making it difficult to produce system designs with guaranteed robust performance. Cyber-physical systems (CPS) is a new multi-disciplinary field aimed at providing a rigorous framework for designing and analyzing these systems, and recent developments in CPS-related fields provide techniques to increase robustness in the design and analysis of complex systems. This seminar brought together researchers from both academia and industry working in hybrid control systems, mechatronics, formal methods, and real-time embedded systems. Participants identified and discussed newly available techniques related to robust design and analysis that could be applied to open issues in the area of CPS and identified open issues and research questions that require collaboration between the communities. This report documents the program and the outcomes of Dagstuhl Seminar 16362 "Robustness in Cyber-Physical Systems".

Seminar September 4–9, 2016 – http://www.dagstuhl.de/16362
1998 ACM Subject Classification C.4 Performance of Systems, C.1.m [Miscellaneous] Hybrid Systems, C.3 Special-Purpose and Application-Based Systems, D.2.4 Software/Program Verification, G.4 Mathematical Software, J.7 Computers in Other Systems
Keywords and phrases aerospace, automotive, cyber-physical systems, fault tolerance, formal verification, real-time and embedded systems, robustness
Digital Object Identifier 10.4230/DagRep.6.9.29

1 Executive Summary

Martin Fränzle
James Kapinski
Pavithra Prabhakar

Overview and Goals of the Seminar

Engineering robustness into systems under development has always been at the heart of good engineering practice, be it robustness against manufacturing tolerances and against variations in purity of construction materials in mechanical engineering, robustness against concentrations of educts in chemical engineering, against parameter variations in the plant model within control engineering, against quantization and measurement noise in signal processing, against faults in computer architecture, against attacks in security engineering, or against unexpected inputs or results in programming. In cyber-physical systems (CPS),

Robustness in Cyber-Physical Systems, *Dagstuhl Reports*, Vol. 6, Issue 9, pp. 29–45
Editors: Martin Fränzle, James Kapinski, and Pavithra Prabhakar

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

all the aforementioned engineering disciplines meet, as the digital networking and embedded control involved in CPS brings many kinds of physical processes into the sphere of human and computer control. This convergence of disciplines has proven extremely fruitful in the past, inspiring profound research on hybrid and distributed control, transferring notions and methods for safety verification from computer science to control theory, transferring proof methods for stability from control theory to computer science, and shedding light on the complex interplay of control objectives and security threats, to name just a few of the many interdisciplinary breakthroughs achieved over the past two decades. Unfortunately, a joint, interdisciplinary approach to robustness remains evasive. While most researchers in the field of CPS concede that unifying notions across the disciplinary borders to reflect the close functional dependencies between heterogeneous components would be of utmost importance, the current state of affairs is a fragmentary coverage by the aforementioned disciplinary notions.

Synergies and research questions

The seminar set out to close the gap in the robustness investigations across the overlapping disciplines under the umbrella of CPS by gathering scientists from the entire spectrum of fields involved in the development of cyber-physical systems and their pertinent design theories. The seminar fostered interdisciplinary research answering the following central questions:

1. What is the rationale behind the plethora of existing notions of robustness and how are they related?
2. What measures have to be taken in a particular design domain (e.g., embedded software design) to be faithful to notions of robustness central to another domain it has functional impact on (e.g., feedback control)?
3. What forms of correctness guarantees are provided by the different notions of robustness and would there be potential for unification or synergy?
4. What design measures have been established by different disciplines for achieving robustness by construction, and how can they be lifted to other disciplines?
5. Where do current notions of robustness or current techniques of system design fall short and can this be alleviated by adopting ideas from related disciplines?

The overarching objective of such research would be to establish trusted engineering approaches incorporating methods for producing cyber-physical system designs

1. that sustain their correctness and performance guarantees even when used in a well-defined vicinity of their nominal operational regimes, and
2. that can be trusted to degrade gracefully even when some of the underlying modeling and analysis assumptions turn out to be false.

To satisfy these design objectives, we require notions of robustness that go well beyond the classical impurities of embedded systems, like sampling, measurement noise, jitter, and machine tolerances, and must draw on concepts of robustness from disparate fields. This seminar identified parallels between related notions of robustness from the many varied domains related to CPS design and bridged the divide between disciplines, with the goal of achieving the above objectives.

Topics of the Seminar

This seminar aimed to identify fundamental similarities and distinctions between various notions of robustness and accompanying design and analysis methods, with the goal of bringing together disparate notions of robustness from multiple academic disciplines and application domains. The following is a brief compendium of the robustness notions and application domains that were addressed in this seminar.

Robustness Notions and Design/Analysis Methods

One goal of this seminar was to identify crosscutting frameworks and design methodologies among the different approaches used to study robustness in the domains of control theory, computer science, and mechanical engineering. We considered the following broad classifications of robustness with the ultimate goal of synergizing the notions and techniques from the various disciplines.

- Input/Output Robustness
- Robustness with respect to system parameters
- Robustness in real-time system implementation
- Robustness due to unpredictable environments
- Robustness to Faults

Application Domains

The applications for the topics addressed in this seminar include cyber-physical systems for which robustness is a vital concern. The following is a partial list of these application domains.

- Automotive
- Aeronautics
- Medical devices
- Robotics
- Smart buildings
- Smart infrastructure

Outcome

We summarize the outcomes of the discussions in the break-out sessions that were conducted by forming subgroups among the participants. The topics referred to different approaches and/or applications in the framework of robustness. One of the topics was about robustness for discrete systems. In this session, the need for defining robustness for these systems was extensively discussed, and one of the most relevant challenges identified was to define appropriate metrics on the state-space relevant to the application. Also some specific robustness issues in the domain of medical devices and automotive systems were identified.

Another discussion was about guaranteeing robust performance from systems based on machine learning. This issue is a difficult task and it is growing in importance as many new safety critical applications, such as self-driving cars, are being designed using machine learning techniques. A challenge is to develop reliable methodologies for certifying or designing for robust performance for systems based on machine learning.

Discussions in a third break-out group were centered around the issue of established engineering means for obtaining robustness by design and how to accommodate these in

rigorous safety cases or formal proofs of correctness. A finding was that most formal models would currently require rather low-level coding of the dynamic behavior of such mechanisms, thereby requiring them to be re-evaluated on each new design rather than exploiting their guaranteed properties to simplify system analysis, which would be in line with their actual impact on engineering processes.

2 Table of Contents

3 Overview of Talks

3.1 Conformance-based robust semantics, and application to anytime control

Houssam Abbas (University of Pennsylvania – Philadelphia, US)

We first describe a Skorokhod-like distance between signals, and generalize the robust semantics of MTL to base them on this new distance. We show that even though the new distance is not a metric, the resulting semantics still satisfy the fundamental properties of (metric-based) robust semantics. In particular, they can be used in a falsification framework. This opens the way to a principled application of robustness-guided falsification to application domains in hybrid systems where the difference between signals might not be adequately captured by the sup norm or other metrics. This new distance was motivated by work in the verification of cardiac devices, where it was found to provide better discrimination between fatal and non-fatal arrhythmias.

We next explore how to use the robust semantics for Anytime control: consider a controller that is being fed noisy state estimates. Can the controller make requests to the estimator, telling it to supply an estimate within a certain time delay, and with a certain error bound? This capability can be used by the controller to save computation power or perform "last-millisecond" aggressive maneuvers. When the control objective is low-level, we present a Model Predictive Control-based solution. We explore how a similar paradigm can be applied to higher-level specifications.

3.2 On Discrete Robustness in Controller Synthesis

Rüdiger Ehlers (Universität Bremen, DE)

Joint work of Rüdiger Ehlers, Ufuk Topcu

A classical approach to CPS control is to first compute a faithful discrete abstraction of the physical environment and to then synthesize a discrete controller that ensures that the specification is satisfied on the discrete abstraction. The approach splits the question of how to obtain robust controllers, i.e., those that can tolerate deviations from the modeled environment conditions whenever possible, into two parts: (1) ensuring robustness of the discrete controller against glitches in the (discrete) abstraction of the environment and (2) making the execution of the continuous actions as robust as possible. We will reconsider the former problem in this talk and study the question if we can infer how the system should behave in case of environment assumption failures from the specification for the nominal operation case. A simple example shows that this is frequently not the case. The example is followed by an outlook on an approach to integrate the system engineer's application knowledge of what constitutes robust behavior into the synthesis process of robust CPS controllers in the future.

3.3 Automatic Test Generation for Autonomous Vehicular Systems

Georgios Fainekos (Arizona State University – Tempe, US)

Joint work of Cumhur Erkan Tuncali, Theodore P. Pavlic
Main reference C. E. Tuncali, T. P. Pavlic, G. Fainekos, "Utilizing S-TaLiRo as an Automatic Test Generation Framework for Autonomous Vehicles," in Proc. of the 19th IEEE Int'l Conf. on Intelligent Transportation Systems (ITCS'16), pp. 1470–1475, IEEE, 2016.
URL http://dx.doi.org/10.1109/ITSC.2016.7795751

Dynamic safety for autonomous vehicular systems is easy to define: avoid collisions at all costs. This definition leads to a natural notion of robustness: keep the distance from all objects of interest as large as possible. Similarly, for passive safety, a system is more robust when the damage to the vehicle is minimized. Even though such notions of robustness may be useful for system design, they are not necessarily useful for the automatic test generation and falsification problems for dynamic safety. Falsification seeks to detect system behaviors that exhibit minimum robustness. Under these metrics, it is easy to produce scenarios where the system under test fails unavoidably and catastrophically. In this work, we define a robustness metric (or, more accurately, a cost function) that combines notions of dynamic and passive safety in order to detect boundary conditions between safe and unsafe behaviors. We demonstrate our results on a simple scenario of autonomous vehicles driving on a multi-lane road.

3.4 When Robustness Comes for Free – Towards Laws of Large Numbers for Ultra-High Integrity Systems

Martin Fränzle (Universität Oldenburg, DE)

Joint work of Martin Fränzle, Sebastian Gerwinn, Ingo Stierand

Statistical physics successfully derives almost sure – i.e., very robust- properties of large ensembles from unpredictable component behavior. Given that cyber-physical systems (CPS) are in fact large ensembles of components, we address the question whether we may expect similar emergent properties of ensembles within CPS and whether these implicitly robustify our systems, giving "robustness for free". We exemplify that effect and demonstrate the underlying mathematics on a single example we very recently have successfully analyzed. It deals with the hard real-time analysis of task systems and is meant to serve as a demonstrator shedding light on the more general applicability of the concept.

Historically in research on real-time systems, hard real-time (in the sense that missing a deadline may have catastrophic effect on the system or its environment) has always been identified with worst-case timing (in the sense of worst-case execution times of tasks, worst-case end-to-end latencies in circuits or reactive systems, etc.). The question, however, is whether this identification is scientifically valid? Given that, e.g., the likelihood of actually encountering the worst-case execution time (WCET) of a single task in a task system already is low (which is why empirical WCET determination is so hard), the probability of simultaneously encountering close to worst-case behavior on most tasks in a set of hundreds of tasks seems to be bound to be astronomically low – probably too low to even worry about. Do we thus really need to care for the sum of the individual tasks' WCETs when computing

the utilization, response time, etc., in the various established schedulability checks? Or would a weaker criterion suffice to establish likelihoods of deadline hits high enough to be acceptable even for extreme integrity systems in highly safety-critical domains?

To address these questions, we set up a formal model facilitating to compute rigorous answers to this question. We therefore reconsider the notion of hard real-time, giving it a stochastic tweak of extremely high confidence rather than sure dead-line hit, and devise a pertinent formal model and analysis method. The reader should note that the question at hand is very different from average-case analysis, which can be pursued with scrutiny by various techniques, among them statistical model-checking (SMC) as a general-purpose tool not requiring any particular theory development. The assurance levels we want to achieve are, however, far beyond its scope, which proves both a burden, as the straightforward techniques like SMC fail, and a virtue, permitting us to set up a powerful approximation theory for those rare events. This theory rigorously proves that the likelihood of a full task system to exceed a certain percentile – say 90%.

3.5 Automatically Robustifying Verified Hybrid Systems in KeYmaera X

Nathan Fulton (Carnegie Mellon University – Pittsburgh, US)

Formal verification of realistic hybrid systems models is an iterative endeavor. Verification efforts typically begin with a simple system model that elides most sources of uncertainty and disturbance. After this relatively simple verification task is completed, the model is robustified against sensing error, actuation uncertainty, plant disturbances, adversarial environments, and other sources of uncertainty or disturbance that arise during testing and simulation. Each new source of uncertainty or disturbance further complicates the model and therefore requires a systematic but none-the-less time-intensive re-verification.

This talk presents early work toward a systematic approach for automatically hardening previously verified hybrid systems against sources of uncertainty and disturbance without requiring re-verification of the robustified system, and discusses an ongoing implementation of this technique in the KeYmaera X theorem prover.

3.6 An algorithmic approach to global asymptotic stability verification of hybrid systems

Miriam García Soto (IMDEA Software – Madrid, ES)

Joint work of Pavithra Prabhakar, Miriam García Soto
Main reference P. Prabhakar, M. García Soto, "An algorithmic approach to global asymptotic stability verification of hybrid systems", in Proc. of the 2016 Int'l Conf. on Embedded Software (EMSOFT'16), pp. 9:1–9:10, ACM, 2016.
URL http://dx.doi.org/10.1145/2968478.2968483

I will present an algorithmic approach to global asymptotic stability (GAS) verification of hybrid systems. Global asymptotic stability is a fundamental property in control system

design which states that small perturbations in the equilibrium point result in only small perturbations in the behaviour of the system, and every execution of the system converges to the equilibrium point. The broad approach is to reduce GAS verification to local asymptotic stability (AS) and region stability (RS) verification. The AS problem is solved by using a quantitative predicate abstraction technique which is also used to compute a stability zone. The RS problem is stated with respect to the stability zone an it is solved by applying an abstraction technique and by performing a termination analysis over it. Positive results of both verification problems result in GAS of the hybrid system. The GAS analysis theory is developed for the case of polyhedral switched systems. The technique is applied to an automatic gearbox model, and provides a GAS proof for this model. Most of the analysis is automated except for certain tasks such as the predicate selection defining the stability zone.

3.7 Automated Checking and Generation of Invariant Sets

Khalil Ghorbal (INRIA – Rennes, FR)

We focus on dynamical systems described by ordinary differential equations with polynomial right-hand side. We investigate two questions of interest for those systems: (i) decision procedures for the invariance of semi-algebraic sets for a given dynamical system, and (ii) the automated generation of invariant algebraic and semi-algebraic sets. We enumerate and theoretically compare previously reported methods as well as the most recent ones. We also empirically assess the practical running performance of such methods on a generic set of benchmarks. The advantages and limitations of such methods will be clearly established thought out the talk.

3.8 Connecting Robust Design with Testing

James Kapinski (Toyota Technical Center – Gardena, US)

Robust design paradigms provide the capability of designing systems that meet performance standards in the presence of parameter variations and disturbances, but they do not guarantee that the deployed system exhibits robust performance. Testing is required to ensure that the system that is ultimately realized displays robust performance. The goal of robust design techniques can therefore be viewed as a means to reduce the amount of testing required to achieve the necessary level of robust performance. This talk argues that artifacts obtained through robust design practices should be used to reduce the effort involved in the test and calibration phases of development. Also, knowledge gained through tests should be used to update the abstractions used in the robust design phase.

3.9 Useful Robustness Notions For Some Industrial Examples

Jens Oehlerking (Robert Bosch GmbH – Stuttgart, DE)

A plethora of robustness notions have been defined in recent years for many model classes and engineering domains. In this talk, three example system from the automotive industry were presented, focusing on useful robustness notions that can be interpreted by engineers. In general, robustness notions tend to be more useful in this context, if they can be traced back to quantities over which the engineer has some form of control. This includes (both physical and non-physical) system parameters, as well as control inputs. In contrast to this, many robustness notions provided by academia focus on quantifying the distance of output signals to desirable or undesirable behavior, leading to robustness metrics that cannot easily be interpreted by an engineer. While such metrics are still very useful (e.g., in the context of optimization based test case generation), it seems that they are not ideal with an engineer in the loop. Therefore, in this talk, a parallel was drawn to approaches for the inversion of dynamical systems, e.g, flatness-based feedforward control. There, the goal is to derive an optimal control input signal given a desired control output signal based on an inverse model. Since it seems that some kind of inverse model is also needed to map robustness notions on system output back onto robustness notions of system inputs or parameters, the question was raised whether this would be a useful research direction.

3.10 Automata-based approach to measuring robustness

Jan Otop (University of Wroclaw, PL)

Joint work of Thomas A. Henzinger, Jan Otop, Roopsha Samanta
Main reference T. A. Henzinger, J. Otop, "Model measuring for discrete and hybrid systems", Nonlinear Analysis: Hybrid Systems, Vol. 23, pp. 166–190, 2017.
URL http://dx.doi.org/10.1016/j.nahs.2016.09.001

Robust systems are the one that continue to work correctly despite of perturbations. The perturbation model is crucial here; we therefore refer to robustness of a system with respect to specific perturbations. Also, it is unlikely that a system is completely immune to all perturbations. This motivates quantitative approach to robustness, where systems are characterized by the level of (specific) perturbations, which they tolerate.

In this talk, I present an automata-based approach to robustness, where perturbations are modeled by weighted automata. The resulting frameworks subsume (some) previously studied notions of robustness, and allow for modelling of a wide range of perturbations, which are additionally graded. Grading perturbations enables us to measure robustness (i.e., establish the level of perturbations safe for the system).

3.11 Robustness for compositional control design

Necmiye Ozay (University of Michigan – Ann Arbor, US)

Joint work of Stanley Smith, Petter Nilsson, Necmiye Ozay

Composing controllers designed individually for interacting subsystems, while preserving the guarantees that each controller provides on each subsystem is a challenging task. In this talk, I will present some of our recent work on using robust control design techniques for compositional design of complex decentralized safety controllers for cyber-physical systems. I will start by introducing some classical qualitative and quantitative notions of robustness in control and estimation. Then, I will present a method for synthesis of controlled invariant sets and associated controllers, that is robust against affine parametric uncertainties in the system matrices. Given a complex system composed of linear parameter varying subsystems, where the system matrices of each subsystem depend (possibly nonlinearly) on the states of the other subsystems, this method can be used for separately designing controllers for subsystems if the uncertainty imposed by a subsystem onto others can be quantified. I will present asymptotically tight techniques for quantification of the uncertainty. Finally, an application of the overall design methodology to vehicle safety systems will be presented. In particular, I will demonstrate how controllers for lane-keeping and adaptive cruise control can be synthesized in a compositional way using the proposed techniques. Our simulations illustrate how these controllers keep their individual safety guarantees when implemented simultaneously, as the theory suggests.

References

1 S. W. Smith, P. Nilsson, and N. Ozay, "Interdependence quantification for compositional control synthesis with an application in vehicle safety systems", Proc. 55th IEEE Conference on Decision and Control (CDC), Las Vegas, NV, December 2016.
2 P. Nilsson and N. Ozay, "Synthesis of separable controlled invariant sets for modular local control design", Proc. American Control Conference (ACC), Boston, MA, July 2016.

3.12 Pre-orders for Reasoning about Stability Properties of Hybrid Systems

Pavithra Prabhakar (Kansas State University – Manhattan, US)

An important class of robustness specifications in control system design is stability. Stability captures the property that small perturbations in the initial state or input lead to only small deviations in the system behavior. We discuss the generalization of stability notions to hybrid systems, and investigate preorders on hybrid systems that preserve stability. The preorders strengthen the classical notions of simulations/bisimulations with uniform continuity conditions that forces preservation of the stability notions.

3.13 Uncertainty handling and robustness analysis of finite precision implementations

Sylvie Putot (Ecole Polytechnique – Palaiseau, FR)

Joint work of Eric Goubault, Sylvie Putot
Main reference E. Goubault, S. Putot, "Robustness analysis of finite precision implementations", in Proc. of the 11th Asian Symp. on Programming Languages and Systems (APLAS'13), LNCS, Vol. 8301, pp. 50–57, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-319-03542-0_4

A desirable property of control systems is robustness to inputs, when small perturbations of the inputs of a system will cause only small perturbations on outputs. This property should be maintained at the implementation level, where close inputs can lead to different execution paths. The problem becomes crucial for finite precision implementations, where any elementary computation is affected by an error. In this context, almost every test is potentially unstable, that is, for a given input, the finite precision and real numbers paths may differ. Still, state-of-the-art error analyses often rely on the stable test hypothesis, yielding unsound error bounds when the conditional block is not robust to uncertainties. We propose an abstract-interpretation based error analysis of finite precision implementations, which is sound in presence of unstable tests, by bounding the discontinuity error for path divergences. This gives a tractable analysis implemented in the FLUCTUAT analyzer.

3.14 Deciding the Undecidable

Stefan Ratschan (The Czech Academy of Sciences – Prague, CZ)

Main reference P. Franek, S. Ratschan, P. Zgliczynski, "Quasi-decidability of a Fragment of the First-Order Theory of Real Numbers", Journal of Autom. Reasoning, 57(2):157–185, 2016.
URL http://dx.doi.org/10.1007/s10817-015-9351-3

Every engineer strives for robustness, simply because models of physical systems have to be robust to be able to work in practice. But robustness has another advantage: it is beneficial for computation. Especially, undecidable problems can become solvable under the assumption of robustness. In the talk, I discussed some results in this direction.

3.15 Towards Robustness for Cyber-Physical Systems

Matthias Rungger (TU München, DE), Sina Caliskan, Rupak Majumdar (MPI-SWS – Kaiserslautern, DE), and Paulo Tabuada (University of California at Los Angeles, US)

Main reference P. Tabuada, S. Y. Caliskan, M. Rungger, R. Majumdar, "Towards Robustness for Cyber-Physical Systems", IEEE Trans. Automat. Contr., 59(12):3151–3163, 2014.
URL http://dx.doi.org/10.1109/TAC.2014.2351632

Robustness as a system property describes the degree to which a system is able to function correctly in the presence of disturbances, i.e., unforeseen or erroneous inputs. In this talk, we present a notion of robustness termed input-output dynamical stability for cyber-physical

systems (CPS) which merges existing notions of robustness for continuous systems and discrete systems. The notion captures two intuitive aims of robustness: bounded disturbances have bounded effects and the consequences of a sporadic disturbance disappear over time. For cyber systems modeled as finite-state transducers, the proposed notion of robustness can be verified in pseudo-polynomial time. The synthesis problem, consisting of designing a controller enforcing robustness, can also be solved in pseudo-polynomial time.

3.16 Robust Cyber-Physical Systems: An utopia within reach

Paulo Tabuada (University of California at Los Angeles, US)

Joint work of Daniel Neider, Paulo Tabuada
Main reference P. Tabuada, D. Neider, "Robust Linear Temporal Logic", in Proc. of the 25th EACSL Annual Conference on Computer Science Logic (CSL'16), LIPIcs, Vol. 62, pp. 10:1–10:21, Schloss Dagstuhl, 2016.
URL http://dx.doi.org/10.4230/LIPIcs.CSL.2016.10

Robustness plays a major role in the analysis and design of engineering systems. Although robust control is a well established area within control theory and fault-tolerant computation is a well established area within computer science, it is surprising that robustness remains a distant mirage for Cyber-Physical Systems. The intricate crochet made of control, computation, and communication yarns is known to be brittle in the sense that "small" software errors or "small" sensing, communication, or actuation noise can lead to unexpected, and often unintended, consequences. In this talk I will build on classical notions of robustness from control theory and computer science to make progress towards the utopia of robust Cyber-Physical Systems.

3.17 Temporal-logic-constrained synthesis and verification without discretization

Ufuk Topcu (University of Texas – Austin, US)

Joint work of Ivan Papusha, Jie Fu, Ufuk Topcu, Richard Murray, Tichakorn Wongpiromsarn, Andrew Lamperski

Can we algorithmically synthesize temporal-logic-constrained controllers for dynamical systems with 50 continuous states? Using conventional methods based on discretization, the answer is 'no'. Even the coarsest discretization would result in intractably large discrete state spaces.

We present a novel approach that avoids explicit discretization in synthesis. We investigate the synthesis of optimal controllers for continuous-time and continuous-state systems under temporal logic specifications. We consider a setting in which the specification can be expressed as a deterministic, finite automaton (the specification automaton) with transition costs, and the optimal system behavior is captured by a cost function that is integrated over time. Specifically, we construct a dynamic programming problem over the product of the underlying continuous-time, continuous-state system and the discrete specification automaton. This dynamic programming formulation relies on the optimal substructure of the additive transition costs over the product of the system and specification automaton. Furthermore,

we propose synthesis algorithms based on approximate dynamic programming for both linear and nonlinear systems under temporal logic constraints. We show that, for linear systems under co-safe temporal logic constraints, this approximate dynamic programming solution reduces to a semidefinite program.

As time allows, we overview a similar approach for the dual problem of verification of dynamical systems against temporal logic specifications. This approach combines automata-based verification and the use of so-called barrier certificates.

References

1 Ivan Papusha, Jie Fu, Ufuk Topcu and Richard Murray. *Automata Theory Meets Approximate Dynamic Programming: Optimal Control with Temporal Logic Constraints*. Conference on Decision and Control, 2016.

2 Tichakorn Wongpiromsarn, Ufuk Topcu and Andrew Lamperski. *Automata theory meets barrier certificates: Temporal logic verification of nonlinear systems*. IEEE Transactions on Automatic Control, http://dx.doi.org/10.1109/TAC.2015.2511722, 2015.

3.18 Robustness in Self-Driving Cars

Eric M. Wolff (nuTonomy – Cambridge, US)

Joint work of nuTonomy team

Self-driving cars are poised to revolutionize transportation, potentially making travel safer, cheaper, and more efficient. Numerous teams have demonstrated autonomous driving on public roads with a safety driver, but there are key technical challenges that must be answered before the safety driver can be removed.

In this talk, I will overview the (public) state-of-the-art in self-driving cars, specifically related to verification and validation. I will introduce different notions of robustness as related to planning, control, perception, and localization, and discuss how careful composition of these subsystems can make the entire system more robust and easier to validate.

3.19 Probabilistic Reachability for Hybrid Systems with Uncertain Parameters

Paolo Zuliani (University of Newcastle, GB) and Fedor Shmarov

Joint work of Fedor Shmarov, Paolo Zuliani
Main reference F. Shmarov, P. Zuliani, "Probabilistic Hybrid Systems Verification via SMT and Monte Carlo Techniques," in Proc. of the 12th Haifa Verification Conf. (HVC'16), LNCS, Vol. 10028, pp. 152–168, Springer, 2016.
URL http://dx.doi.org/10.1007/978-3-319-49052-6_10

Hybrid systems are a framework much used for modelling cyber-physical systems, and are finding more application in other areas, such as systems biology and systems medicine. Reachability is a key verification analysis: in this talk I will focus on bounded reachability, i.e., in a finite number of steps (or jumps). If a hybrid system contains random parameters, then reachability amounts to computing a probability; if the system also features uncertain

(nondeterministic) parameters, then reachability generalises to finding enclosures for reachability probabilities. In this talk I will survey our two approaches to probabilistic bounded reachability. One is fully rigorous – and comes high computational complexity – and one is a mixture of a rigorous and a statistical approach, thereby yielding better scalability by trading absolute guarantees with statistical guarantees.

Participants

- Houssam Abbas
University of Pennsylvania – Philadelphia, US
- Paul Bogdan
USC – Los Angeles, US
- Alexandre Donzé
University of California – Berkeley, US
- Rüdiger Ehlers
Universität Bremen, DE
- Georgios Fainekos
Arizona State University – Tempe, US
- Martin Fränzle
Universität Oldenburg, DE
- Nathan Fulton
Carnegie Mellon University – Pittsburgh, US
- Miriam García Soto
IMDEA Software – Madrid, ES
- Khalil Ghorbal
INRIA – Rennes, FR
- James Kapinski
Toyota Technical Center – Gardena, US
- Scott C. Livingston
Washington D.C., US
- Sarah M. Loos
Google Research, US
- Rupak Majumdar
MPI-SWS – Kaiserslautern, DE
- Jens Oehlerking
Robert Bosch GmbH – Stuttgart, DE
- Jan Otop
University of Wroclaw, PL
- Necmiye Ozay
University of Michigan – Ann Arbor, US
- Pavithra Prabhakar
Kansas State University – Manhattan, US
- Sylvie Putot
Ecole Polytechnique – Palaiseau, FR
- Stefan Ratschan
The Czech Academy of Sciences – Prague, CZ
- Matthias Rungger
TU München, DE
- Paulo Tabuada
University of California at Los Angeles, US
- Ufuk Topcu
University of Texas – Austin, US
- Eric M. Wolff
nuTonomy – Cambridge, US
- Bai Xue
Universität Oldenburg, DE
- Paolo Zuliani
University of Newcastle, GB

Report from Dagstuhl Seminar 16371

Public-Key Cryptography

Edited by

Marc Fischlin[1], Alexander May[2], David Pointcheval[3], and Tal Rabin[4]

1 **TU Darmstadt, DE,** marc.fischlin@cryptoplexity.de
2 **Ruhr-Universität Bochum, DE,** alex.may@ruhr-uni-bochum.de
3 **ENS – Paris, FR,** david.pointcheval@ens.fr
4 **IBM Thomas J. Research Center – Yorktown Heights, US,** talr@us.ibm.com

Abstract

This report documents the program and results of Dagstuhl seminar 16731 "Public-Key Cryptography" which took place September 11–16, 2016. The goal of the seminar was to bring together different subareas from public-key cryptography and to promote research among these areas.

Seminar September 11–16, 2016 – http://www.dagstuhl.de/16371
1998 ACM Subject Classification D.4.6 Security and Protection, E.3 Data Encryption
Keywords and phrases cryptanalysis, encryption, homomorphic encryption, key exchange, obfuscation, signatures
Digital Object Identifier 10.4230/DagRep.6.9.46
Edited in cooperation with Sven Schäge

1 Summary

Marc Fischlin

Cryptography has turned out to be an invaluable tool for protecting the confidentiality and integrity of digital data. At the same time, cryptography does not yet provide satisfying solutions to all practical scenarios and threats. To accomplish appropriate protection of the data, cryptography needs to address several challenges.

Cryptography has always been a prominent theme within the Dagstuhl Seminar series, with the first meeting about cryptography held in 1993, and subsequent seminars on this topic about every 5 years. In 2007 and 2012 a seminar for the subarea of "Symmetric Cryptography" has been added, inciting us to coin the seminar here "Public-Key Cryptography" for sake of distinction. The public-key branch has been held for the second time, after the first event in 2011.

The seminar brought together 27 scientists in the area of public-key cryptography, including three student researchers who were invited by Dagstuhl to pick a seminar to participate in. The participants came from all over the world, including countries like the US, Great Britain, Israel, France, or Japan. Among the affiliations, Germany lead the number with 9 participants, followed by the US and France with 6 each. The program contained 21 talks, each of 25 to 60 minutes, and a panel discussion about the uneasiness with the current state of our reviewing system, with a free afternoon on Wednesday for social activities and the afternoon on Thursday for collaborations. Before the seminar, we asked the participants to present very recent and ongoing work which, ideally, should not have been published or

accepted to publication yet. Most of the participants followed our suggestion and to a large extend the presentations covered topics which have not even been submitted at the time.

The topics of the talks represented the diversity of public-key cryptography. The goal of the seminar was to bring together three challenge areas in cryptography, namely, cryptanalysis and foundations (investigating and evaluating new primitives), optimization (making solutions more efficient), and deployment (designing real-world protocols). As envisioned, the seminar thus has a good mixture of talks from these areas. There were also suggestions to try to co-locate future events of the seminar with other security-related events at Dagstuhl to foster even broader interdisciplinary research. Discussions during and after the talks were lively. It seems as if the goal of stimulating collaborations among these areas has been met. The discussion about the reviewing system has led to some hands-on practices which could be deployed to improve the quality of reviews. This includes incentives such as"Best Reviewer Awards" and teaching students about proper reviewing.

2 Table of Contents

3 Overview of Talks

3.1 Diverse Vector Spaces and Zero-Knowledge

Fabrice Benhamouda (IBM Thomas J. Watson Research Center – Yorktown Heights, US)

Joint work of Michel Abdallah, Fabrice Benhamouda, David Pointcheval
Main reference M. Abdallah, F. Benhamouda, D. Pointcheval, "Disjunctions for Hash Proof Systems: New Constructions and Applications," in Proc. of the 34th Annual Int'l Conf. on the Theory and Applications of Cryptographic Techniques – Advances in Cryptology (EUROCRYPT'15) – Part II, LNCS, Vol. 9057, pp. 69–100, Springer, 2015; pre-print available at IACR.
URL http://dx.doi.org/10.1007/978-3-662-46803-6_3
URL https://eprint.iacr.org/2014/483

We first present hash proof systems or smooth projective hash functions (SPHFs), which were introduced by Cramer and Shoup in 2002 to explain the construction of the Cramer-Shoup IND-CCA encryption scheme and which later found numerous other applications. We then introduce diverse vector spaces as a tool to construct SPHFs. Finally, we illustrate this tool on simple examples and show applications to zero-knowledge primitives.

3.2 What Else is Revealed by Order-Revealing Encryption?

David Cash (Rutgers University, US)

Joint work of F. Betül Durak, Thomas M. DuBuisson, David Cash
Main reference F. Betül Durak, Thomas M. DuBuisson, David Cash, "What Else is Revealed by Order-Revealing Encryption?," in Proc. of the 2016 ACM SIGSAC Conf. on Computer and Communications Security (CCS'16), pp. 1155–1166, ACM, 2016; pre-print available at IACR.
URL http://dx.doi.org/10.1145/2976749.2978379
URL http://eprint.iacr.org/2016/786

The security of order-revealing encryption (ORE) has been unclear since its invention. Dataset characteristics for which ORE is especially insecure have been identified, such as small message spaces and low-entropy distributions. On the other hand, properties like one-wayness on uniformly-distributed datasets have been proved for ORE constructions.

This work shows that more plaintext information can be extracted from ORE ciphertexts than was previously thought. We identify two issues: First, we show that when multiple columns of correlated data are encrypted with ORE, attacks can use the encrypted columns together to reveal more information than prior attacks could extract from the columns individually. Second, we apply known attacks, and develop new attacks, to show that the leakage of concrete ORE schemes on non-uniform data leads to more accurate plaintext recovery than is suggested by the security theorems which only dealt with uniform inputs.

3.3 Comparison between Subfield and Straightforward Attacks on NTRU

Pierre-Alain Fouque (University of Rennes, FR)

Joint work of Paul Kirchner, Pierre-Alain Fouque

Recently in two independent papers, Albrecht, Bai and Ducas and Cheon, Jeong and Lee presented two very similar attacks, that allow to break NTRU with larger parameters and GGH Multinear Map without zero encodings. They proposed an algorithm for recovering the NTRU secret key given the public key which apply for large NTRU modulus, in particular to Fully Homomorphic Encryption schemes based on NTRU. Hopefully, these attacks do not endanger the security of the NTRUE NCRYPT scheme, but shed new light on the hardness of this problem. The basic idea of both attacks relies on decreasing the dimension of the NTRU lattice using the multiplication matrix by the norm (resp. trace) of the public key in some subfield instead of the public key itself. Since the dimension of the subfield is smaller, the dimension of the lattice decreases, and lattice reduction algorithm will perform better. Here, we revisit the attacks on NTRU and propose another variant that is simpler and outperforms both of these attacks in practice. It allows to break several concrete instances of YASHE, a NTRU-based FHE scheme, but it is not as efficient as the hybrid method of Howgrave-Graham on concrete parameters of NTRU. Instead of using the norm and trace, we propose to use the multiplication by the public key in some subring and show that this choice leads to better attacks. We can then show that for power of two cyclotomic fields, the time complexity is polynomial. Finally, we show that, under heuristics, straightforward lattice reduction is even more efficient, allowing to extend this result to fields without non-trivial subfields, such as NTRU Prime. We insist that the improvement on the analysis applies even for relatively small modulus ; though if the secret is sparse, it may not be the fastest attack. We also derive a tight estimation of security for (Ring-)LWE and NTRU assumptions. when $q = 2^{\Omega(\sqrt{n \log \log n})}$.

3.4 Advances in building Non-Malleable Commitments

Vipul Goyal (Microsoft Research India – Bangalore, IN)

Joint work of Vipul Goyal, Omkant Pandey, Silas Richelson

A central challenge in the design of secure systems is to defend against man-in-the-middle attacks, where an adversary can arbitrarily tamper with the messages exchanged by two parties over a communication channel. Starting with the early nineties, an important research goal in cryptography has been to build "non malleable" cryptographic protocols that are resilient to such attacks.

A very basic non-malleable primitive which is widely used in cryptography is what is known as non-malleable commitment schemes. In this talk, I will describe a recent result which constructs non-malleable commitments in the minimal number of rounds (and almost minimal complexity assumptions). In some sense, this culminates a two-decade long research quest of getting non-malleable commitments in the minimal number of rounds.

3.5 Fair Coin Flipping: Tighter Analysis and the Many-Party Case

Iftach Haitner (Tel Aviv University, IL)

Joint work of Niv Buchbinder, Iftach Haitner, Levi Nissan, Eliad Tsfadia

In a multi-party fair coin-flipping protocol, the parties output a common (close to) unbiased bit, even when some corrupted parties try to bias the output. In this work we focus on the case of dishonest majority, i.e. at least half of the parties can be corrupted. Cleve (STOC 1986) has shown that in any m-round coin-flipping protocol the corrupted parties can bias the honest parties' common output bit by $1/m$. For more than two decades the best known coin-flipping protocols against dishonest majority was the protocol of Awerbuch, Blum, Chor, Goldwasser, and Micali [Manuscript 85], who presented a t-party, m-round protocol of bias $t/\sqrt{m}$. This was changed by the breakthrough result of Moran, Naor and Segev (TCC 2009), who constructed an m-round, 2-party coin-flipping protocol with optimal bias of $1/m$. Recently, Haitner and Tsafadia (STOC 14) constructed an m-round, three-party coin-flipping protocol with bias $O(log^3(m)/m)$. Still for the case of more than three parties, the best known protocol remains the $\Theta(t/\sqrt{m})$-bias protocol of Awerbuch et al.

We make a step towards eliminating the above gap, presenting a t-party, m-round coin-flipping protocol, with bias $O(\frac{t*2^t*\sqrt{\log m}}{m^{1/2+1/(2^t-1-2)}})$. This improves upon the $\Theta(t/\sqrt{m})$-bias protocol of Awerbuch et al. for any $t < 1/2 * log(log(m))$, and in particular for $t \in O(1)$, this yields an $1/m^{1/2+\Theta(1)}$-bias protocol. For the three-party case, this yields an $O(\sqrt{logm}/m)$-bias protocol, improving over the $O(log^3 m/m)$-bias protocol of Haitner and Tsafadia. Our protocol generalizes that of Haitner and Tsafadia, by presenting an appropriate "defense protocols" for the remaining parties to interact in, in the case that some parties abort or caught cheating (Haitner and Tsafadia only presented a two-party defense protocol, which limits their final protocol to handle three parties).

We analyze our new protocols by presenting a new paradigm for analyzing fairness of coin-flipping protocols. We map the set of adversarial strategies that try to bias the honest parties outcome in the protocol to the set of the feasible solutions of a linear program. The gain each strategy achieves is the value of the corresponding solution. We then bound the optimal value of the linear program by constructing a feasible solution to its dual.

3.6 Kurosawa-Desmedt Meets Tight Security

Dennis Hofheinz (KIT – Karlsruher Institut für Technologie, DE)

Joint work of Romain Gay, Dennis Hofheinz, Lisa Kohl

At EUROCRYPT 2016, Gay et al. presented the first pairing-free public-key encryption (PKE) scheme with a tight security reduction to a standard assumption. Their scheme is competitive in efficiency with state-of-the art PKE schemes and has very compact ciphertexts (of three group elements), but suffers from a large public key (of about 200 group elements).

In this work, we present an improved pairing-free PKE scheme with a tight security reduction to the Decisional Diffie-Hellman assumption, small ciphertexts (of three group elements), *and* small public keys (of six group elements). Compared to the work of Gay et al.,

our scheme thus has a considerably smaller public key and comparable other characteristics, although our encryption and decryption algorithms are somewhat less efficient.

Technically, our scheme borrows ideas both from the work of Gay et al. and from a recent work of Hofheinz (eprint, 2016). The core technical novelty of our work is an efficient and compact designated-verifier proof system for an OR-like language. We show that adding such an OR-proof to the ciphertext of the state-of-the-art PKE scheme from Kurosawa and Desmedt enables a tight security reduction.

3.7 Schnorr Signatures in the Multi-User Setting

Eike Kiltz (Ruhr-Universität Bochum, DE)

Joint work of Eike Kiltz, Daniel Masny, Jiaxin Pan
Main reference E. Kiltz, D. Masny, J. Pan, "Optimal Security Proofs for Signatures from Identification Schemes," in Proc. of the 36th Annual Int'l Cryptology Conf. – Advances in Cryptology (CRYPTO'16) – Part II, LNCS, Vol. 9815, pp. 33–61, Springer, 2016.
URL http://dx.doi.org/10.1007/978-3-662-53008-5_2

A theorem by Galbraith, Malone-Lee, and Smart (GMLS) from 2002 showed that, for Schnorr signatures, single-user security tightly implies multi-user security. Recently, Bernstein pointed to an error in the above theorem and promoted a key-prefixing variant of Schnorr signatures for which he proved a tight implication from single to multi-user security. Even worse, he identified an "apparently insurmountable obstacle to the claimed [GMLS] theorem". This paper shows that, without key prefixing, single-user security of Schnorr signatures tightly implies multi-user security of the same scheme. Our result has slightly stronger requirements than the GLML theorem: we either require the random oracle model or strong single user security of Schnorr signatures.

3.8 Computational Arithmetic Secret Sharing

Alexander Koch (KIT – Karlsruher Institut für Technologie, DE)

Secret sharing schemes allow for sharing a secret message so that it can be correctly reconstructed in the presence of enough of its shares, but with the property that nothing can be learned about its content if too few of the shares have been obtained. Homomorphic schemes exhibit the additional property that it is possible to calculate on the shares to obtain a share of the sums and products of secrets, yielding a plethora of applications including secure multiparty computation (MPC). To reduce the size of the generated shares in a secret sharing scheme,"computational" variants have been developed which guarantee secrecy for illegitimate access to the secret only against computationally restricted adversaries. While these schemes are much more size-efficient, they usually have the disadvantage of not being homomorphic. We give the first computational secret sharing scheme on the basis of multi-key fully homomorphic encryption, that combines the advantages of both worlds.

3.9 Practical LPN Cryptanalysis

Alexander May (Ruhr-Universität Bochum, DE)

Joint work of Andre Esser, Robert Kübler, Alexander May

We present memory-efficient algorithms for LPN, both classically and quantumly. We also show first experiments for solving LPN instances up to dimension 250 with error parameter 1/8.

3.10 Concurrently Composable Security With Shielded Super-polynomial Simulators

Jörn Müller-Quade (KIT – Karlsruher Institut für Technologie, DE)

Joint work of Brandon Broadnax, Nico Döttling, Gunnar Hartung, Jörn Müller-Quade, Matthias Nagel
Main reference B. Broadnax, N. Döttling, G. Hartung, J. Müller-Quade, M. Nagel, "Concurrently Composable Security With Shielded Super-polynomial Simulators", Cryptology ePrint Archive, Report 2016/1043, 2016.
URL http://eprint.iacr.org/2016/1043

We propose a new framework for concurrently composable security that relaxes the security notion of UC security. As in previous frameworks, our notion is based on the idea of providing the simulator with super-polynomial resources. However, in our new framework simulators are only given *restricted access* to the results computed in super-polynomial time. This is done by modeling the super-polynomial resource as a stateful oracle that may directly interact with a functionality without the simulator seeing the communication. We call these oracles "shielded oracles".

Our notion is fully compatible with the UC framework, i. e., protocols proven secure in the UC framework remain secure in our framework. Furthermore, our notion lies strictly between SPS and Angel-based security, while being closed under protocol composition.

Shielding away super-polynomial resources allows us to apply new proof techniques where we can replace super-polynomial entities by indistinguishable polynomially bounded entities. This allows us to construct secure protocols in the plain model using weaker primitives than in previous composable frameworks involving simulators with super-poly resources. In particular, we only use non-adaptive-CCA-secure commitments as a building block in our constructions. As a feasibility result, we present a constant-round general MPC protocol in the plain model based on standard assumptions that is secure in our framework.

3.11 Lattice Enumeration Revisited

Phong Q. Nguyen (Inria and CNRS/JFLI, FR, and University of Tokyo, JP)

Lattice enumeration is arguably the simplest method to solve exact lattice problems. Though it does not have the best asymptotical time complexity, it has been used in the largest lattice records, notably NTRU challenges, Darmstadt's lattice challenges and SVP challenges. In this talk, we revisit lattice enumeration with pruning techniques.

3.12 Overcoming Hellman's Time/Memory Trade Offs with Applications to Proofs of Space

Krzysztof Pietrzak (IST Austria – Klosterneuburg, AT)

Joint work of Bram Cohen, Danylo Khilko, Hamza Abusalah, Joel Alwen, Krzysztof Pietrzak, Leonid Reyzin

Hellman showed that any permutation over a domain of size N can be inverted in time T by an algorithm whose description is of size S for any S, T which satisfy $N < O(S \cdot T)$ (e.g. $S = T \approx N^{1/2}$), for general functions a weaker attack $N^3 < O(S^3 \cdot T)$ (e.g. $S = T \approx N^{3/4}$) exists.

The best lower bounds are of the form $N > \tilde{\Omega}(S \cdot T)$ and hold for random permutations and functions.

Motivated by the application to proofs of space (PoSpace), we construct functions for which we can prove much better lower bounds of the form $N^k > \tilde{\Omega}(S^k \cdot T)$ (for any constant k). Our construction does not contradict the existing attacks, as these attacks require that the function to be inverted can be efficiently computed in forward direction. For the application to PoSpace it is sufficient that the entire function table can be computed in time quasilinear in N.

The simplest function that beats the existing bound is build from a random function $g : [N] \times [N] \to [N]$ and a random permutations $f, f' : [N] \to [N]$ and is defined as $h(x) = g(x, x')$ where $f(x) = f(x') + 1$ (instead of $+1$ one can use any other bijection without fixpoints). For this function we prove a lower bound of $N^2 < O(S^2 \cdot T)$. Note that h cannot be efficiently evaluated on input x as one has on find $x' = f^{-1}(f(x) - 1)$, but its function table can be computed in time $O(N)$ by first computing the function table for f^{-1}.

3.13 Integer Commitments

David Pointcheval (ENS – Paris, FR)

Joint work of Geoffroy Couteau, Thomas Peters
Main reference G. Couteau, T. Peters, D. Pointcheval, "Removing the Strong RSA Assumption from Arguments over the Integers", Cryptology ePrint Archive, Report 2016/128, 2016.
URL http://eprint.iacr.org/2016/128

Committing integers and proving relations between them is an essential ingredient in many cryptographic protocols. Among them, range proofs have shown to be fundamental. They consist in proving that a committed integer lies in a public interval. By the way, it can also be seen as a particular case of the more general Diophantine relations: for the committed vector of integers $\vec{x}$, there exists a vector of integers $\vec{w}$ such that $P(\vec{x}, \vec{w}) = 0$, where P is a polynomial.

In this talk, we revisit the security strength of the statistically hiding commitment scheme over the integers due to Damgård-Fujisaki, and the zero-knowledge proofs of knowledge of openings.

First, we show how to remove the Strong RSA assumption and replace it by the standard RSA assumption in the security proofs. This improvement naturally extends to generalized commitments and more complex proofs without modifying the original protocols.

Second, we design an interactive technique turning commitment scheme over the integers into commitment scheme modulo a prime p. Still under the RSA assumption, this results in more efficient proofs of relations between the committed values. Our methods thus improve upon existing proof systems regarding Diophantine relations both in terms of performance and security.

We illustrate that with more efficient range proofs under the sole RSA assumption.

3.14 Securing Public Key Encryption in the Presence of Bad Randomness

Jacob Schuldt (AIST – Tsukuba, JP)

Joint work of Takahiro Matsuda, Kenny Paterson, Jacob Schuldt, Dale Sibborn, Hoeteck Wee

In this talk, we firstly motivate the need for encryption secure in the presence of bad randomness, and revisit the notion of related randomness security by Paterson, Schuldt, and Sibborn, as well as some of the known constructions of related randomness secure encryption. We then highlight an inherent limitation of the related randomness security notion: if the family of related randomness functions is sufficiently rich to express the encryption function of the considered scheme, then security cannot be achieved. This might help explain why the previous standard model constructions only achieve security for polynomial function families.

To address this limitation, we propose a new notion, related refreshable randomness security, which captures that an adversary has limited time to attack a system before new entropy is added. In this setting, we construct an encryption scheme which remains secure in the standard model for arbitrary function families of size 2^p (where p is polynomial in the security parameter) that satisfy certain collision-resistant and output-unpredictability properties. This captures a rich class of functions, which includes, as a special case, circuits of polynomial size.

3.15 On the Impossibility of Tight Cryptographic Reductions

Sven Schäge (Ruhr-Universität Bochum, DE)

Joint work of Christoph Bader, Tibor Jager, Yong Li, Sven Schäge
Main reference C. Bader, T. Jager, Y. Li, S. Schäge, "On the Impossibility of Tight Cryptographic Reductions," Proc. of the 35th Annual Int'l Conf. on the Theory and Applications of Cryptographic Techniques – Advances in Cryptology (EUROCRYPT'16) – Part II, LNCS, Vol. 9666, pp. 273–304, Springer, 2016.
URL http://dx.doi.org/10.1007/978-3-662-49896-5_10

The existence of tight reductions in cryptographic security proofs is an important question, motivated by the theoretical search for cryptosystems whose security guarantees are truly independent of adversarial behavior and the practical necessity of concrete security bounds for the theoretically-sound selection of cryptographic parameters. At Eurocrypt 2002, Coron described a meta-reduction technique that allows to prove the impossibility of tight reductions for certain digital signature schemes. This seminal result has found many further interesting applications. However, due to a technical subtlety in the argument, the applicability of this

technique beyond digital signatures in the single-user setting has turned out to be rather limited.

We describe a new meta-reduction technique for proving such impossibility results, which improves on known ones in several ways. First, it enables interesting novel applications. This includes a formal proof that for certain cryptographic primitives (including public-key encryption/key encapsulation mechanisms and digital signatures), the security loss incurred when the primitive is transferred from an idealized single-user setting to the more realistic multi-user setting is impossible to avoid, and a lower tightness bound for non-interactive key exchange protocols. Second, the technique allows to rule out tight reductions from a very general class of non-interactive complexity assumptions. Third, the provided bounds are quantitatively and qualitatively better, yet simpler, than the bounds derived from Coron's technique and its extensions.

3.16 Android Security using Static Analysis Techniques

Suzanna Schmeelk (Columbia University – New York, US)

Joint work of Suzanna Schmeelk, Alfred Aho, Junfeng Yang
URL https://www.cs.columbia.edu/~schmeelk/publications.html

Static analysis is a traditional technique for software transformation and analysis. It has also become a means to detect cyber security vulnerabilities and malware and recently has been extended to the mobile-computing arena for security-related analyses. This talk examines over fifty recent security papers that are published in top conferences, journals and technical reports and characterizes the current research. The papers were selected based on either their high citings by other top research or they introduced either a novel analysis technique or a novel security issue analysis. Our research systematically constructs a static analysis landscape by charting and characterizing analysis strengths and limitations in both accuracy and security threats. It identifies two types of static analysis motivations which affect the soundness of an analysis methodology: (1) techniques for analyzing software for vulnerabilities and (2) techniques used to examine applications for malware, which may lead to malware mitigation. We analyze techniques and tools for effort-level required use by security analysists and connect the reported static analysis motivations to both Mitre's attack taxonomy as well as Mitre's vulnerability taxonomy to aid completeness. Our findings include identifying vulnerabilities which are not being systematically researched, identifying best practices for developers and characterizing technique usability metrics for integrating the analysis into a security analysis process.

3.17 The OPTLS Protocol and TLS 1.3

Hoeteck Wee (ENS – Paris, FR)

Joint work of Hugo Krawczyk, Hoeteck Wee
Main reference H. Krawczyk, H. Wee, "The OPTLS Protocol and TLS 1.3", in Proc. of the IEEE Europ. Symp. on Security and Privacy (EuroS&P'16), pp. 81–96, IEEE, 2016; pre-print available at IACR.
URL http://dx.doi.org/10.1109/EuroSP.2016.18
URL http://eprint.iacr.org/2015/978

We present the OPTLS key-exchange protocol, its design, rationale and cryptographic analysis. OPTLS design has been motivated by the ongoing work in the TLS working group of the IETF for specifying TLS 1.3, the next-generation TLS protocol. The latter effort is intended to revamp the security of TLS that has been shown inadequate in many instances as well as to add new security and functional features. The main additions that influence the cryptographic design of TLS 1.3 (hence also of OPTLS) are a new "0-RTT requirement" (0-RTT stands for "zero round trip time") to allow clients that have a previously retrieved or cached public key of the server to send protected data already in the first flow of the protocol; making forward secrecy (PFS) a mandatory requirement; and moving to elliptic curves as the main cryptographic basis for the protocol (for performance and security reasons). Accommodating these requirements calls for moving away from the traditional RSA-centric design of TLS in favor of a protocol based on Diffie-Hellman techniques. OPTLS offers a simple design framework that supports all the above requirements with a uniform and modular logic that helps in the specification, analysis, performance optimization, and future maintenance of the protocol. An earlier (draft) specification of TLS 1.3 built upon the OPTLS framework as a basis for the cryptographic core of the handshake protocol, adapting the different modes of OPTLS and its HKDF-based key derivation to the TLS 1.3 context.

Participants

- Adekunle Oluseyi Afolabi
University of Kuopio, FI
- Fabrice Benhamouda
IBM Thomas J. Watson Research Center – Yorktown Heights, US
- Johannes A. Buchmann
TU Darmstadt, DE
- David Cash
Rutgers University, US
- Pooya Farshim
ENS – Paris, FR
- Marc Fischlin
TU Darmstadt, DE
- Pierre-Alain Fouque
University of Rennes, FR
- Vipul Goyal
Microsoft Research India – Bangalore, IN
- Iftach Haitner
Tel Aviv University, IL
- Dennis Hofheinz
KIT – Karlsruher Institut für Technologie, DE
- Antoine Joux
CNRS and University Pierre & Marie Curie – Paris, FR
- Eike Kiltz
Ruhr-Universität Bochum, DE
- Alexander Koch
KIT – Karlsruher Institut für Technologie, DE
- Tal Malkin
Columbia Univ. – New York, US
- Alexander May
Ruhr-Universität Bochum, DE
- Jörn Müller-Quade
KIT – Karlsruher Institut für Technologie, DE
- Phong Q. Nguyen
Inria and CNRS/JFLI, FR, and University of Tokyo, JP
- Kenneth G. Paterson
Royal Holloway University of London, GB
- Krzysztof Pietrzak
IST Austria – Klosterneuburg, AT
- David Pointcheval
ENS – Paris, FR
- Tal Rabin
IBM Thomas J. Watson Research Center – Yorktown Heights, US
- Sven Schäge
Ruhr-Universität Bochum, DE
- Suzanna Schmeelk
Columbia Univ. – New York, US
- Dominique Schröder
Univ. Erlangen-Nürnberg, DE
- Jacob Schuldt
AIST – Tsukuba, JP
- Vinod Vaikuntanathan
MIT – Cambridge, US
- Hoeteck Wee
ENS – Paris, FR

Report from Dagstuhl Seminar 16372

Uncertainty Quantification and High Performance Computing

Edited by

Vincent Heuveline[1], Michael Schick[2], Clayton Webster[3], and Peter Zaspel[4]

1 **HITS & Universität Heidelberg, DE,** vincent.heuveline@h-its.org
2 **Robert Bosch GmbH – Stuttgart, DE,** michael.schick3@de.bosch.com
3 **Oak Ridge National Laboratory, US,** webstercg@ornl.gov
4 **HITS & Universität Heidelberg, DE,** peter.zaspel@uni-heidelberg.de

Abstract

High performance computing is a key technology to solve large-scale real-world simulation problems on parallel computers. Simulations for a fixed, deterministic set of parameters are current state of the art. However, there is a growing demand in methods to appropriately cope with uncertainties in those input parameters. This is addressed in the developing research field of uncertainty quantification. Here, Monte-Carlo methods are easy to parallelize and thus fit well for parallel computing. However, their weak approximation capabilities lead to inaccurate results. The Dagstuhl Seminar 16372 "Uncertainty Quantification and High Performance Computing" brought together experts in the fields of uncertainty quantification and high performance computing. Discussions on the latest numerical techniques beyond pure Monte-Carlo and with strong approximation capabilities were fostered. This has been put in context of real-world problems on parallel computers.

Seminar September 11–16, 2016 – http://www.dagstuhl.de/16372
1998 ACM Subject Classification D.1.3 Concurrent Programming, G.1.2 Approximation, G.3 Probability and Statistics
Keywords and phrases high performance computing, parallelization, stochastic modeling, uncertainty quantification
Digital Object Identifier 10.4230/DagRep.6.9.59

1 Executive Summary

Vincent Heuveline
Michael Schick
Clayton Webster
Peter Zaspel

Topics

Uncertainty quantification (UQ) aims at approximating measures for the impact of uncertainties in e.g. simulation parameters or simulation domains. By this way, it is of great importance for both academic research and industrial development. In uncertainty quantification, one distinguishes between classical forward uncertainty propagation and more involved inference, optimization or control problems under uncertainties. Forward uncertainty

Uncertainty Quantification and High Performance Computing, *Dagstuhl Reports*, Vol. 6, Issue 9, pp. 59–73
Editors: Vincent Heuveline, Michael Schick, Clayton Webster, and Peter Zaspel

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

propagation is concerned with deterministic numerical models for e.g. engineering problems, in which parts of the input data (domain, parameters, ...) might be affected by uncertainties, i.e. they have a random nature. Randomness is usually characterized by random fields that replace the originally deterministic inputs. In Bayesian inference, parameters of a system shall be derived for given measurements. Since the measurements are assumed to be affected by some (stochastic) error, this inference approach tries to derive probabilities under which a given parameter leads to the observed measurements. In some sense, Bayesian inference complements classical inverse problems in a stochastic sense. Other fields of interest for a similar uncertainty analysis are optimization and control.

High performance computing (HPC) is an interdisciplinary research field in computer science, mathematics and engineering. Its aim is to develop hardware, algorithmic approaches and software to solve (usually) mathematically formulated problems on large clusters of interconnected computers. The dominant part of the involved research is done in parallel computing. From a hardware perspective, HPC or parallel computing requires to develop computing technologies that can e.g. solve several problems at the same time at high performance and low power. Moreover, hardware developments in HPC often aim at improving network communication technologies, which are necessary to let a (potentially) large set of computers solve a single problem in a distributed way. From an algorithmic perspective, methods known from numerical mathematics and data processing are adapted such that they can run in a distributed way on different computers. Here, a key notion is (parallel) scalability which describes the ability to improve the performance or throughput of a given method by increasing the number of used computers. Most algorithmic developments shall improve this scalability for numerical methods. Research in software aims at defining appropriate programming models for parallel algorithms, providing efficient management layers for the underlying hardware and implementing the proposed parallel algorithms in real software.

Challenges

In UQ, (partial) differential equations with random data are approximately solved by either intrusive or non-intrusive methods. An intrusive technique simultaneously discretizes stochastic and physical space with the classical example of stochastic Galerkin approaches. This method delivers favorable properties such as small errors with fewer number of equations and potentially small overall run-time. To achieve that, it requires to re-discretize and re-implement existing deterministic PDE solvers. On the other hand, non-intrusive techniques (e.g. (quasi-)Monte Carlo, multi-level Monte Carlo, stochastic collocation, ...) reuse existing solvers / simulation tools and generate a series of deterministic solutions which are used to approximate stochastic moments. It is thereby possible to perform uncertainty quantification analysis even for very complex large-scale applications for which a re-implementation of existing solvers is no option. The non-intrusive approach is connected to a rather extreme computational effort, with at least hundreds, thousands or even more deterministic problems that have to be solved. While a single real-world forward uncertainty propagation problem is already extremely computational intensive, even on a larger parallel computer, inference, optimization and control under uncertainties often go beyond the limits of currently available parallel computers.

In HPC, we have to distinguish methods that are intrinsically (often also called embarrassingly) parallel and those that have to exchange data to compute a result. That

is, embarrassingly parallel algorithms are able to independently compute on completely decoupled parts of a given problem. A prominent example in UQ are Monte-Carlo-type methods. The other extreme are approaches that require to exchange a lot of data in order to solve a given problem. Here, prominent examples are adaptive and multi-level methods in general and stochastic Galerkin methods. Both method types tend to have excellent approximation properties, but require a considerable effort in parallel algorithms to be scalable on parallel computers. Scalability considerations might become even more important on the next generation of the largest parallel computers, which are expected to be available at the beginning of the next decade. These parallel Exascale computers will be able to process on the exaFLOP level, thus they will be able to issue 10 18 floating-point instructions within a second. Technological limitations in chip production will force computing centers to install systems with a parallel processor count which is by orders of magnitude higher than in current systems. Current parallel algorithms might not be prepared for this next step.

The Dagstuhl Seminar on "Uncertainty Quantification and High Performance Computing", brought together experts from UQ and HPC to discuss some of the following challenging questions:

- How can real-world forward uncertainty problems or even inference, control and optimization under uncertainties be made tractable by high performance computing?
- What types of numerical uncertainty quantification approaches are able to scale on current or future parallel computers, without sticking to pure Monte Carlo methods?
- Might adaptivity, model reduction or similar techniques improve existing uncertainty quantification approaches, without breaking their parallel performance?
- Can we efficiently use Exascale computing for large-scale uncertainty quantification problems without being affected by performance, scalability and resilience problems?
- Does current research in uncertainty quantification fit the needs of industrial users? Would industrial users be willing and able to use HPC systems to solve uncertainty quantification problems?

Seminar outcome

Several presentations covered Bayesion inference / inversion (Ghattas, Marzouk, Najm, Peters), where seismology is an extremely computationally expensive problem that can only be solved by the largest parallel computers (Ghattas). While the parallelization is crucial, the numerical methods have to be adapted as well, such that fast convergence is achieved (Ghattas, Marzouk, Peters). The very computationally intensive optimization under uncertainties (Benner) becomes tractable by the use of tensor approximation methods (Benner, Osedelets). Tensor approximation methods as well as hierarchical matrices (Börm, Zaspel) are optimal complexity numerical methods for a series of applications in UQ. However their large-scale parallelization is still subject to research.

A series of talks considered mesh-free approximation methods (Rieger, Teckentrup, Zaspel) with examples in Gaussian process regression (Teckentrup) and kernel-based methods. It was possible to see that these methods have provable error bounds (Rieger, Teckentrup) and can be scaled on parallel computers (Rieger, Zaspel). Moreover these methods even fit well for inference (Teckentrup). Sparse grid techniques were considered as example for classical approximation methods for higher-dimensional problems (Stoyanov, Peters, Harbrecht, Pflüger). Here, recent developments in adaptivity and optimal convergence were discussed. Sparse grid techniques are usually considered in a non-intrusive setting such

that parallel scalability is often guaranteed. Compressed sensing promises to reduce the amount of simulations in a non-intrusive framework (Dexter). Quasi-Monte Carlo methods are under investigation for optimal convergence (Nuyens). The latter methods are of high interest for excellent parallel scalability on parallel computers due to the full decoupling of all deterministic PDE solves while keeping convergence orders beyond classical Monte Carlo methods.

Adaptivity leads to strongly improved approximations using the same amount of deterministic PDE solutions (Pflüger, Stoyanov, Webster, ...). However, a clear statement on how to parallelize adaptive schemes in an efficient way is still subject to research. The general class of multi-level schemes was also under investigation (Dodwell, Zhang), including but not being limited to multi-level Monte-Carlo and multi-level reduced basis approaches. These methods show excellent convergence properties. However their efficient and scalable parallelization is part of intensive studies, as well.

Performance considerations in the field of HPC (including future parallel computers) have been discussed (Heuveline, Legrand). Performance predictability is necessary to understand scaling behavior of parallel codes on future machines (Legrand). Parallel scalability of (elliptic) stochastic PDEs by domain decomposition has been discussed by LeMaître. His approach allows to increase parallel scalability and might show hints towards resilience.

Industrial applications were considered for the company Bosch (Schick), where intrusive and non-intrusive approaches are under investigation. High performance computing is still subject to discussion in this industrial context. One of the key applications, which is expected to become an industrial-like application, is UQ in medical engineering (Heuveline). Once introduced into the daily work cycle at hospitals, it will soon become a driving technology for our health.

Perspectives

Based on the survey and personal feedback from the invitees, the general consensus is that there is a high interest in deepening the discussions at the border of UQ and HPC. While some answers to the above questions could be given, there is still a lot more to learn, to discuss and to develop. A general wish is therefore to have similar meetings in the future.

Acknowledgements. The organizers would like to express their gratitude to all participants of the Seminar. Special thanks go to the Schloss Dagstuhl team for its extremely friendly support during the preparation phase and for the warm welcome at Schloss Dagstuhl.

2 Table of Contents

3 Overview of Talks

3.1 Optimization of Random Navier-Stokes Equations

Peter Benner (MPI – Magdeburg, DE)

We discuss the optimization and optimal control of flow problems described by the unsteady, incompressible Navier-Stokes equations. Randomness is introduced by modeling the uncertainty in the dynamic viscosity as a random variable. Using a stochastic discretization of the optimality system leads to a large-scale nonlinear system of equations in saddle point form. Nonlinearity is treated with a Picard-type iteration in which linear saddle point systems have to be solved in each iteration step. Using data compression based on separation of variables and the tensor train (TT) format, we show how these large-scale indefinite and nonsymmetric systems that typically have 10^8–10^{11} unknowns can be solved without the use of HPC technology. The key observation is that the unknown and the data can be well approximated in a new block TT format that reduces complexity by several orders of magnitude. We illustrate our findings by numerical examples.

3.2 Hierarchical tensor approximation

Steffen Börm (Universität Kiel, DE)

Joint work of Steffen Börm, Dirk Boysen, Isabelle Greff

We consider the computation of two-point correlations of the stochastic partial differential equation

$$-\Delta u(x,\omega) = f(x,\omega),$$

where x is a point in a domain D and ω is an element of a probability space. Following a result by Schwab and Todor, the two-point correlations C_u satisfy the equation

$$\Delta_x \Delta_y C_u(x,y) = C_f(x,y),$$

where C_f denotes the two-point correlations of the right-hand side. Since this is an equation in $D \times D$, the computational cost of standard discretization schemes is fairly high even if D is only a two-dimensional domain.

We propose an alternative approach: an analysis by Pentenrieder and Schwab indicates that C_u is smooth in large parts of the domain $D \times D$, so it is possible to approximate the solution by an *hp* finite element method. In order to avoid having to construct a locally refined mesh for the four- or even six-dimensional domain $D \times D$, we employ a hierarchical partition of unity in combination with suitable tensor-product functions. We introduce a recursive algorithm for constructing the sparsity pattern of the resulting system matrix. This algorithm also suggests a technique for obtaining the matrix coefficients based only on the system matrix of the original partial differential equation by using suitable inter-grid transfer operators.

3.3 Towards UQ + HPC for Bayesian Inversion, with Application to Global Seismology

Omar Ghattas (University of Texas at Austin, US)

Joint work of Tan Bui, Carsten Burstedde, Pearl Flath, Omar Ghattas, James Martin, Georg Stadler, Hari Sundar, Lucas Wilcox

Inverse problems governed by acoustic, elastic, or electromagnetic wave propagation – in which we seek to reconstruct the unknown shape of a scatterer, or the unknown properties of a medium, from observations of waves that are scattered by the shape or medium – play an important role in a number of engineered or natural systems. Our goal is to address the quantification of uncertainty in the solution of the inverse problem by casting the inverse problem as one in Bayesian inference. This provides a systematic and coherent treatment of uncertainties in all components of the inverse problem, from observations to prior knowledge to the wave propagation model, yielding the uncertainty in the inferred medium/shape in a systematic and consistent manner. Unfortunately, state-of-the-art MCMC methods for characterizing the solution of Bayesian inverse problems are prohibitive when the forward problem is expensive (as in our 100–1000 wavelength target problems) and a high-dimensional parametrization is employed to describe the unknown medium (as in our target problems involving infinite-dimensional medium/shape fields, which result in millions of parameters when discretized).

The Hessian operator of the negative log posterior plays an important role in the efficient solution of Bayesian inverse problems. When the parameter-to-observable map is linearized at the MAP point (and the prior and noise are Gaussian), the posterior is a Gaussian with the inverse Hessian as its covariance operator. More generally, this geometry-aware Gaussian approximation can be used within a proposal to accelerate MCMC methods for sampling non-Gaussian posteriors, such as in the so-called stochastic Newton, Riemannian manifold, or DILI MCMC methods.

The Hessian is often the sum of a compact operator (the data misfit) and an elliptic differential operator (the inverse prior), and this invites a low-rank approximation of the (prior-preconditioned) data misfit term, leading to an effective reduction in dimensionality (often several orders of magnitude).

Here we show that the following combination of conditions leads to a class of methods whose cost (measured in forward/adjoint PDE solves) scales independent of the parameter and data dimensions and number of processor cores:

- the prior-preconditioned data misfit Hessian is compact with mesh and data independent dominant spectrum (typical of ill-posed inverse problems)
- dominant spectrum is captured in O(r) matvecs with Hessian (we use randomized SVD for the low rank approximation)
- Hessian-vector products are computed matrix-free using second-order adjoint-based methods (amounts to 2 linearized PDE solves per matvec)
- fast O(n) elliptic solvers used for prior operator applications (we use hybrid geometric/algebraic multigrid to handle heterogeneous/anisotropic priors)
- the forward and adjoint PDE solves scale well with number of cores

The cost to construct the Laplace approximation of the posterior (or the local Gaussian at every MCMC iteration when the posterior is sufficiently non-Gaussian) is overwhelmingly

dominated by O(r) linearized forward/adjoint PDE solves (to construct the low-rank approximation). Everything else is negligible linear algebra. So when the PDE forward/adjoint solver scales well, one achieves a scalable UQ method. For mildly non-Gaussian posteriors, we evaluate the Hessian and its inverse at the maximum a posteriori point and reuse it during the MCMC iterations.

For strongly non-Gaussian posteriors, the inverse Hessian formally has to be computed repeatedly, which is intractable for large-scale, high-dimensional problems, even if the number of (linearized) forward solves is independent of the parameter and data dimensions. The challenge is to find better representations of the Hessian beyond low rank (e.g. H-matrix-based) or else to derive effective preconditioners (e.g. based on its symbol).

We present applications to a Bayesian inverse problem in global seismology with up to one million earth model parameters and 630 million state variables, on up to 100,000 processor cores.

3.4 Solution of free boundary problems in the presence of geometric uncertainties

Helmut Harbrecht (Universität Basel, CH), Marc Dambrine, Michael Peters, and Benedicte Puig

The solution of Bernoulli's exterior free boundary problem is considered in case of an interior boundary which is random. Two ways are introduced to define the expectation and the deviation of the resulting annular domain. To compare both approaches, some analytical examples for a circular interior boundary are studied. Moreover, numerical experiments are performed for more general geometrical configurations. In order to numerically approximate the expectation and the deviation, a sampling method is proposed like the (quasi-) Monte Carlo quadrature. The free boundary is determined for each sample by the trial method which is a fixed-point like iteration.

References

1 H. Harbrecht and M. Peters. *Solution of free boundary problems in the presence of geometric uncertainties.* Preprint 2015-02, Mathematisches Institut, Universität Basel, Switzerland, 2015 (to appear in Radon Series on Computational and Applied Mathematics, de Gruyter).
2 M. Dambrine, H. Harbrecht, M. Peters, and B. Puig. *On Bernoulli's free boundary problem with a random boundary.* Manuscript, 2016.

3.5 Uncertainty Quantification and High Performance Computing: Quid?

Vincent Heuveline (HITS & Universität Heidelberg)

The increasing demand on reliable results in scientific computing makes the quantification of uncertainties in mathematical models a crucial task. Including Uncertainty Quantification

to scientific computing leads for many applications to a shift of paradigm from purely deterministic problems to the stochastic models. In addition, the development of new technologies in high performance computing enables to consider new numerical methods in order to solve the challenging problems arising in Uncertainty Quantification. The talk adresses the interface between Uncertainty Quantification and High Performance Computing with a main emphasize in:

- intrusive methods in Uncertainty Quantification for systems of partial differential equations (PDEs);
- efficient accelerator and preconditioning technologies to be used on large-scale supercomputing clusters;
- open-source software-development for making the implementations accessible for the worldwide research community.

Applications in medical engineering are presented. A blood pump scenario where the inflow boundary condition, viscosity and the rotation speed are modeled as uncertain parameter is depicted. It shows up both the potential of high performance computing for uncertainty quantification but also still existing numerical challenges for real world applications.

3.6 Performance Prediction of HPC Applications: The SimGrid Project

Arnaud Legrand (INRIA – Grenoble, FR)

Simulation of HPC applications. Parallel platforms have progressively become more and more heterogeneous and complicated. After a quick presentation of typical recent supercomputers, I have presented how task-based programming and dynamic runtimes allow to efficiently exploit such architecture. Yet, evaluating performance of such complex systems is particularly challenging. I have thus presented our recent work on StarPU/SimGrid, a custom simulator that can be used to predict the performance of task-based applications running on top of StarPU to exploit hybrid (CPU+GPU) architectures. We have demonstrated the faithfulness of StarPU/SimGrid for both modern dense and sparse linear algebra solvers.

References

1 Luka Stanisic, Samuel Thibault, Arnaud Legrand, Brice Videau, Jean-François Méhaut. *Faithful Performance Prediction of a Dynamic Task-Based Runtime System for Heterogeneous Multi-Core Architectures*. Concurrency and Computation: Practice and Experience, Wiley, 2015, pp. 16.

2 Luka Stanisic, Emmanuel Agullo, Alfredo Buttari, Abdou Guermouche, Arnaud Legrand, et al.. *Fast and Accurate Simulation of Multithreaded Sparse Linear Algebra Solvers*. The 21st IEEE International Conference on Parallel and Distributed Systems, Dec. 2015, Melbourne, Australia.

3 Henri Casanova, Arnaud Giersch, Arnaud Legrand, Martin Quinson, Frédéric Suter. *Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms*. Journal of Parallel and Distributed Computing, Elsevier, 2014, 74 (10), pp. 2899–2917.

3.7 Quasi-Monte Carlo methods for elliptic PDEs with random coefficients

Dirk Nuyens (KU Leuven, BE)

Joint work of I. Graham, Frances Y. Kuo, Dirk Nuyens, Rob Scheichl, Ian H. Sloan

I first discuss the current theory of getting dimension independent convergence in approximating high-dimensional and infinite-dimensional integrals. This is done by using weighted reproducing kernel Hilbert spaces. For several quasi-Monte Carlo methods we know function spaces and weights for which we have optimal convergence independent of the number of dimensions. Then I apply this theory to a parametrised PDE where we balance the dimension truncation error, FEM error and quadrature/cubature error. For log-normal random fields we obtain dimension-independent convergence of N^{-1} using randomly shifted lattice rules.

3.8 Bayesian Inversion for Electrical Impedance Tomography

Michael Peters (Universität Basel, CH)

Joint work of Robert Gantner, Helmut Harbrecht, Michael Peters, Markus Siebenmorgen

In this talk, we consider a Bayesian approach towards Electrical Impedance Tomography, where we are interested in computing moments, in particular the expectation, of the contour of an unknown inclusion, given noisy current measurements at the surface. By casting the forward problem into the framework of elliptic diffusion problems on random domains, we solve a suitably parametrized version by means of the domain mapping method. This straightforwardly yields parametric regularity results for the system response, which we exploit to conduct a rigorous analysis of the posterior measure, facilitating the application of sophisticated quadrature methods for the approximation of moments of quantities of interest. As an example of such a quadrature method, we consider an anisotropic sparse grid quadrature. To solve the forward problem numerically, we employ a fast boundary integral solver. Numerical examples are provided to illustrate the presented approach and validate the theoretical findings.

References

1 R. N. Gantner, M. D. Peters. *Higher Order Quasi-Monte Carlo for Baysian Shape Inversion.* Preprint 2016-18, Mathematisches Institut, Universität Basel, Switzerland, 2016.
2 A.-L. Haji-Ali, H. Harbrecht, M. Peters, and M. Siebenmorgen. *Novel results for the anisotropic sparse quadrature and their impact on random diffusion problems.* Preprint 2015-27, Mathematisches Institut, Universität Basel, Switzerland, 2015.
3 H. Harbrecht, M. Peters, and M. Siebenmorgen. *Analysis of the domain mapping method for elliptic diffusion problems on random domains.* Numer. Math., 134(4):823–856, 2016.

3.9 From Data to Uncertainty: Efficient Data-Driven Adaptive Sparse Grids for UQ

Dirk Pflüger (Universität Stuttgart, DE)

Joint work of Franzelin, Fabian; Jakeman, John; Pfander, David
Main reference F. Franzelin, D. Pflüger, "From Data to Uncertainty: An Efficient Integrated Data-Driven Sparse Grid Approach to Propagate Uncertainty", in Sparse Grids and Applications – Stuttgart 2014, LNCSE, Vol. 109, pp. 29–49, Springer, 2016.
URL http://dx.doi.org/10.1007/978-3-319-28262-6_2

We consider non-intrusive stochastic collocation for uncertainty quantification, as our applications require us to treat the underlying simulation code as a black box. We propose spatially adaptive sparse grids for both the estimation of the stochastic densities and the stochastic collocation.

With sparse grids, the numerical discretization is still possible in higher-dimensional settings, and the integrated sparse grid approach leads to fast and efficient algorithms and implementations. This allows us to start with data that is provided by measurements and to combine the estimated densities with the model function's surrogate without introducing additional sampling or approximation errors. Bayesian inference and Bayesian updating allow us to incorporate observations and to adaptively refine the surrogate based on the posterior.

Efficient and scalable algorithms for the evaluation of the surrogate function are available, which can achieve close-to-peak performance even on hybrid hardware.

3.10 Kernel methods for large scale data analysis problems arising in UQ

Christian Rieger (Universität Bonn, DE)

Many problems in uncertainty quantification (UQ) are modeled via parametric partial differential equations (PDEs). Here, the parameters often stem from a given high–dimensional space. A typical reconstruction process consists of three steps. In the first step, one has to solve the parametric PDE for a given set of parameter values. The second step is to compute some derived quantity of interest (QoI) such as a mean of the solution of the PDE for a fixed parameter. Hence, one obtains point-evaluations from a function directly mapping from the parameter space to the real numbers describing the QoI as function of the parameter. Both steps involve some numerical procedure and hence introduce a numerical error to the data. As a third step, one is often only interested in approximatively reconstructing the QoI as function from the parameter space, in order to evaluate this function for new parameter values.

In this talk, we focus on the third step. We make use of the fact that the function mapping the parameter space to the QoI is typically a smooth function of the parameter, see [1]. We present different regularization techniques which aim at saving numerical costs and solving the approximation problem up to the numerical evaluation error level stemming from the first tow steps above. To this end, we propose an error balancing strategy where we compare the numerical evaluation error of the quantity interest and the approximation error which

stems from the fact that we use only finitely many data values. Such a balancing requires an a priori error analysis in order to determine accuracy for the numerical evaluation which is needed in the first two steps. We present an error analysis based on sampling inequalities, see [2]. For the approximation we use reproducing kernels of certain problem-adapted Hilbert space, see [2]. For recent numerical examples using kernel based-methods, see also [3].

This talk is based on joint works with M. Griebel (Bonn), T. Hangelbroek (Hawaii), F. Narcowich (Texas A&M), J. Ward (Texas A&M), and P. Zaspel (Heidelberg).

References

1 A. Cohen, R. DeVore, and C. Schwab. *Analytic regularity and polynomial approximation of parametric and stochastic elliptic pde's.* Analysis and Applications, 09(01):11–47, 2011.
2 M. Griebel and C. Rieger. *Reproducing kernel Hilbert spaces for parametric partial differential equations.* 2015. To appear in SIAM/ASA Journal on Uncertainty Quantification, DOI: 10.1137/15M1026870, also available as INS Preprint No. 1511.
3 P. Zaspel. *Parallel RBF Kernel-Based Stochastic Collocation for Large-Scale Random PDEs.* Dissertation, Institut für Numerische Simulation, Universität Bonn, 2015.

3.11 A Dynamically Adaptive Sparse Grids Method for Quasi-Optimal Interpolation of Multidimensional Functions

Miroslav Stoyanov (Oak Ridge National Laboratory, US) and Clayton Webster (Oak Ridge National Laboratory, US)

In this work we develop a dynamically adaptive sparse grids (SG) method for quasi-optimal interpolation of multidimensional analytic functions defined over a product of one dimensional bounded domains. The goal of such approach is to construct an interpolant in space that corresponds to the "best M-terms" based on sharp a priori estimate of polynomial coefficients. In the past, SG methods have been successful in achieving this, with a traditional construction that relies on the solution to a Knapsack problem: only the most profitable hierarchical surpluses are added to the SG. However, this approach requires additional sharp estimates related to the size of the analytic region and the norm of the interpolation operator, i.e., the Lebesgue constant. Instead, we present an iterative SG procedure that adaptively refines an estimate of the region and accounts for the effects of the Lebesgue constant. Our approach does not require any a priori knowledge of the analyticity or operator norm, is easily generalized to both affine and non-affine analytic functions, and can be applied to sparse grids built from one dimensional rules with arbitrary growth of the number of nodes. In several numerical examples, we utilize our dynamically adaptive SG to interpolate quantities of interest related to the solutions of parametrized elliptic and hyperbolic PDEs, and compare the performance of our quasi-optimal interpolant to several alternative SG schemes.

References

1 M. Stoyanov, C. Webster, *A Dynamically Adaptive Sparse Grid Method for Quasi-Optimal Interpolation of Multidimensional Analytic Functions*, Computers & Mathematics with Applications, Vol. 71, Num. 11, pp. 2449–2465, 2016.

3.12 Gaussian process regression in Bayesian inverse problems

Aretha Teckentrup (University of Warwick – Coventry, GB) and Andrew Stuart

Main reference A. M. Stuart, A. L. Teckentrup, "Posterior Consistency for Gaussian Process Approximations of Bayesian Posterior Distributions", arXiv:1603.02004v2 [math.NA], 2016.
URL https://arxiv.org/abs/1603.02004v2

A major challenge in the application of sampling methods to large scale inverse problems, is the high computational cost associated with solving the forward model for a given set of input parameters. To overcome this difficulty, we consider using a surrogate model that approximates the solution of the forward model at a much lower computational cost. We focus in particular on Gaussian process emulators, and analyse the error in the posterior distribution resulting from this approximation.

3.13 Scalable hierarchical methods on many-core hardware – Fast matrix approximations in kernel-based collocation

Peter Zaspel (HITS & Universität Heidelberg)

It is well-known that future parallel hardware architectures will have a constantly growing number of parallel processing units. Nowadays, many-core processors (GPUs, Xeon Phi) give a first insight into the degree of parallelism that we expect to see in the future. However, it is usually said that the extreme parallelism can only be effectively used, if we apply it to "simple" algorithms. On the other hand, current optimal methods for approximation in the field of uncertainty quantification (multi-level / multi-index Monte Carlo, hierarchical matrices, ...) use complex hierarchical / tree constructions to achieve optimal complexities and approximation results. Seemingly, we have two contradicting development directions: (1) simple, very parallel algorithms; (2) complex, optimal algorithms.

This presentation shall shed light on problems and opportunities we face and have on current many-core processors if we use them to execute hierarchical algorithms. We base our discussion on our recent work in the field of radial basis function (RBF) kernel-based stochastic collocation. This non-intrusive approximation method combines high-order algebraic or even exponential convergence rates of spectral (sparse) tensor-product methods with optimal pre-asymptotic convergence of kriging and the profound stochastic framework of Gaussian process regression. Our recent applications for this approach were (elliptic) model problems and incompressible two-phase flows.

One important part of the kernel-based stochastic collocation is the solution of large to huge dense linear systems with Vandermonde-type matrices. This presentation will discuss the efficient parallel and optimal-complexity solution of these kind of linear systems by iterative solvers and fast matrix-approximations by H-matrices on many-core hardware.

Current limitations and opportunities will be highlighted.

3.14 A multilevel reduced-basis method for parameterized partial differential equations

Guannan Zhang (Oak Ridge National Laboratory, US)

Joint work of Miroslav Stoyanov and Clayton Webster

An important approximation scheme for alleviating the overall computational complexity of solving parameterized PDEs is known as multilevel methods, which have been successfully used in the Monte Carlo and collocation setting. In this effort, we propose to improve the multilevel methods with the use of reduced-basis (RB) techniques for constructing the spatial-temporal model hierarchy of PDEs. Instead of approximating the solution manifold of the PDE, the key ingredient is to build approximate manifolds of first-order differences of PDE solutions on consecutive levels. To this end, we utilize a hierarchical finite element (FE) framework to formulate an easy-to-solve variational FE system for the first-order differences. Moreover, by deriving a posteriori error estimates for the RB solutions, we also intend to develop a greedy-type adaptive strategy in order to construct a good set of snapshots. The main advantage of our approach lies in the fact that the manifold of the first-order differences becomes progressively linear as the physical level increases. Thus, much fewer expensive snapshots are required to achieve a prescribed accuracy, resulting in significant reduction of the offline computational cost of greedy algorithms. Furthermore, our approach combines the advantages of both multilevel Monte Carlo and multilevel collocation methods, in the sense that it can generate snapshots anywhere in the parameter domain but also features fast convergence.

Participants

- Peter Benner
MPI – Magdeburg, DE
- Steffen Börm
Universität Kiel, DE
- Nick Dexter
University of Tennessee – Knoxville, US
- Tim Dodwell
University of Exeter, GB
- Omar Ghattas
University of Texas at Austin, US
- Helmut Harbrecht
Universität Basel, CH
- Vincent Heuveline
HITS &
Universität Heidelberg, DE
- Olivier Le Maitre
LIMSI – Orsay, FR
- Arnaud Legrand
INRIA – Grenoble, FR
- Youssef M. Marzouk
MIT – Cambridge, US
- Habib Najm
Sandia Nat. Labs – Livermore, US
- Dirk Nuyens
KU Leuven, BE
- Ivan Oseledets
Skoltech – Moscow, RU
- Michael Peters
Universität Basel, CH
- Dirk Pflüger
Universität Stuttgart, DE
- Christian Rieger
Universität Bonn, DE
- Michael Schick
Robert Bosch GmbH – Stuttgart, DE
- Miroslav Stoyanov
Oak Ridge National Lab., US
- Aretha Teckentrup
University of Warwick – Coventry, GB
- Clayton Webster
Oak Ridge National Lab., US
- Peter Zaspel
HITS &
Universität Heidelberg, DE
- Guannan Zhang
Oak Ridge National Lab., US

Report from Dagstuhl Seminar 16381

SAT and Interactions

Edited by

Olaf Beyersdorff[1], Nadia Creignou[2], Uwe Egly[3], and Heribert Vollmer[4]

1 **University of Leeds, GB,** o.beyersdorff@leeds.ac.uk
2 **Aix-Marseille University, FR,** nadia.creignou@lif.univ-mrs.fr
3 **TU Wien, AT,** uwe@kr.tuwien.ac.at
4 **Leibniz Universität Hannover, DE,** vollmer@thi.uni-hannover.de

Abstract

This report documents the programme and outcomes of Dagstuhl Seminar 16381 "SAT and Interactions". The seminar brought together researchers from different areas from theoretical computer science involved with various aspects of satisfiability. A key objective of the seminar has been to initiate or consolidate discussions among the different groups for a fresh attack on one of the most important problems in theoretical computer science and mathematics.

Seminar September 18–23, 2016 – http://www.dagstuhl.de/16381
1998 ACM Subject Classification E.1 Data Structures, F.2 Analysis of Algorithms and Problem Complexity, G.2.1 Combinatorics
Keywords and phrases Combinatorics, Computational Complexity, P vs. NP, Proof Complexity, Quantified Boolean formulas, SAT-solvers, satisfiability problem
Digital Object Identifier 10.4230/DagRep.6.9.74
Edited in cooperation with Joshua Blinkhorn

1 Executive Summary

Olaf Beyersdorff
Nadia Creignou
Uwe Egly
Heribert Vollmer

Brief Introduction to the Topic

Propositional satisfiability (or Boolean satisfiability) is the problem of determining whether the variables of a Boolean formula can be assigned truth values in such a way as to make the formula true. This satisfiability problem, SAT for short, stands at the crossroads of logic, graph theory, computer science, computer engineering and computational physics. Indeed, many problems originating from one of these fields typically have multiple translations to satisfiability. Unsurprisingly, SAT is of central importance in various areas of computer science including algorithmics, verification, planning, hardware design and artificial intelligence. It can express a wide range of combinatorial problems as well as many real-world ones.

SAT is very significant from a theoretical point of view. Since the Cook-Levin theorem, which identified SAT as the first **NP**-complete problem, it has become a reference for an enormous variety of complexity statements. The most prominent one is the question "is

P equal to **NP**?" Proving that SAT is not in **P** would answer this question negatively. Restrictions and generalizations of the propositional satisfiability problem play a similar rôle in the examination of other complexity classes and relations among them. In particular, quantified versions of SAT (QSAT, in which Boolean variables are universally or existentially quantified) as well as variants of SAT in which some notion of minimality is involved, provide prototypical complete problems for every level of the polynomial hierarchy.

During the past three decades, an impressive array of diverse techniques from mathematical fields, such as propositional and first-order logic, model theory, Boolean function theory, complexity, combinatorics and probability, has contributed to a better understanding of the SAT problem. Although significant progress has been made on several fronts, most of the central questions remain unsolved so far.

One of the main aims of the Dagstuhl seminar was to bring together researchers from different areas of activity in SAT so that they can communicate state-of-the-art advances and embark on a systematic interaction that will enhance the synergy between the different areas.

Concluding Remarks and Future Plans

The organizers regard the seminar as a great success. Bringing together researchers from different areas of theoretical computer science fostered valuable interactions and led to fruitful discussions. Feedback from the participants was very positive as well. Many attendants expressed their wish for a continuation.

Finally, the organizers wish to express their gratitude toward the Scientific Directorate of the Center for its support of this seminar, and hope to be able to continue this series of seminars on *SAT and Interactions* in the future.

2 Table of Contents

3 Organization of the Seminar and Activities

The seminar brought together 39 researchers with complementary expertise from different areas of theoretical computer science and mathematics, such as logic, complexity theory, algorithms and proof complexity. The participants consisted of both senior and junior researchers, including a number of postdocs and a few advanced graduate students.

Participants were invited to present their work and to communicate state-of-the-art advances. Twenty-three talks of various lengths took place over the five days of the seminar. Introductory and tutorial talks of 60 minutes, introducing one particular aspect of the satisfiability problem, were scheduled to open the first four days of the seminar. The rest of the days were filled mostly with shorter talks picking up the topic of the morning talk. The organizers considered it important to leave ample free time for discussion.

In this way, the following topics evolved:

1. Proof complexity
   - Sam Buss: Satisfiability Testing and Proof Complexity (Tutorial)
   - Barnaby Martin: Resolution and the Binary Encoding of Weak Pigeonhole Principles
   - Jakob Nordström: Supercritical Space-Width Trade-offs for Resolution
   - Jan Johannsen: On Linear Resolution – an Update
   - Ilario Bonacina: Strong Size Lower bounds in Regular Resolution via Games
2. Quantified Boolean Formulas: Solvers and Proof Complexity
   - Florian Lonsing: QBF Solving (Tutorial)
   - Marijn Heule: Practical Proof Systems for SAT and QBF
   - Meena Mahajan: QBF Proof Complexity (Tutorial)
   - Joshua Blinkhorn: On Soundness of QBF Calculi Parameterized by Dependency Schemes
   - Anil Shukla: Understanding Cutting Planes for QBFs
   - Leroy Chew: A Class of Hard Formulas for QBF Resolution
3. Exact Algorithms for SAT
   - Rahul Santhanam: Exact Algorithms for SAT – an Overview (Tutorial)
   - Dominik Scheder: The PPSZ Algorithm: Making Hertli's Analysis Simpler and 3-SAT Faster
   - Uwe Schöning: Classroom Analysis of Random Walk Algorithm for 3-SAT and Practical Extension to ProbSAT
   - Victor Lagerkvist: Partial Polymorphisms and the Time Complexity of SAT Problems
4. Knowledge Compilation
   - Stefan Mengel: An Introduction to Knowledge Compilation (Tutorial)
   - Florent Capelli: Compilation of CNF-formulas: Lower and Upper Bounds

There were additionally a few shorter talks covering further topics related to satisfiability.

- Jacobo Torán: Isomorphism of Solution Graphs
- Arne Meier, Irena Schindler: Approaching Backdoors in Two Non-Classical Logics
- Christoph Wintersteiger: Lifting SAT to Richer Theories: Bit-vectors, Finite Bases, and Theory Combination
- Oliver Kullmann: Look-ahead for Solving Hard SAT Problems
- Miki Hermann: Minimal Distance of Propositional Models
- John Franco: Adding Unsafe Constraints to Improve Satisfiability Performance (Redux)

Thursday afternoon was closed with an open problem session (see later in this report).

Wednesday afternoon was devoted to the usual hike. The day ended with a musical event that was highly appreciated by the seminar participants. The programme can be found in this report.

The above classification of topics and talks is necessarily rough, as several talks crossed the boundaries between these areas, in keeping with the theme of the seminar. The broad scope of the talks extended even to areas not anticipated by the organizers, such as dependence logic. The seminar thus achieved its aim of bringing together researchers from various related communities to share state-of-the-art research.

4 Overview of Talks

4.1 On Soundness in QBF Calculi Parameterized by Dependency Schemes

Joshua Blinkhorn (University of Leeds, GB)

Joint work of Olaf Beyersdorff, Joshua Blinkhorn
Main reference O. Beyersdorff, J. Blinkhorn, "Dependency Schemes in QBF Calculi: Semantics and Soundness", in Proc. of the 22nd Int'l Conf. Principles and Practice of Constraint Programming (CP'16), LNCS, Vol. 9892, pp. 96–112, Springer, 2016.
URL http://dx.doi.org/10.1007/978-3-319-44953-1_7

In the talk, we consider the parameterization of QBF resolution calculi by dependency schemes. One of the main problems in this area is to understand for which dependency schemes the resulting calculi are sound. It is known that a property called *full exhibition* is sufficient for soundness in Q-resolution [2]. We demonstrate that this approach generalizes to the dependency versions of all CDCL-based QBF calculi. Moreover, we show that the most important schemes in the literature possess this property; in particular, the reflexive resolution path dependency scheme is fully exhibited.

The talk also presents some new work, exposing similarities between the two currently disparate fields of QBF dependency schemes and dependency quantified Boolean formulas (DQBF). In particular, using results from [1] we show that the DQBF interpretation of dependency schemes leads to a complete characterisation of soundness for expansion-based QBF calculi. The new interpretation also provides a fresh insight for Q-resolution. We show that the phenomenon of incompleteness in the DQBF calculi, observed by [3], is directly related to the characterization of soundness for the dependency QBF systems.

References

1 Beyersdorff, O., Chew, L,. Schmidt, R. A., Suda, M.: Lifting QBF Resolution Calculi to DQBF. International Conference on Theory and Applications of Satisfiability Testing (SAT). pp. 490–499 (2016).
2 Slivovsky, F.: Structure in #SAT and QBF. Ph. D. Thesis, Vienna University of Technology (2015).
3 Balabanov, V., Chiang, H. K., Jiang, J. R.: Henkin quantifiers and Boolean formulae: A certification perspective of DQBF. Theoretical Computer Science 523, pp. 86–100 (2014).

4.2 Strong Size Lower Bounds in Regular Resolution via Games

Ilario Bonacina (KTH Royal Institute of Technology – Stockholm, SE)

Joint work of Ilario Bonacina, Navid Talebanfard
Main reference I. Bonacina, N. Talebanfard, "Strong ETH and Resolution via Games and the Multiplicity of Strategies", Algorithmica, pp. 1–13, Springer, 2016.
URL http://dx.doi.org/10.1007/s00453-016-0228-6

The Strong Exponential Time Hypothesis (SETH) says that solving the SAT problem on formulas that are k-CNFs in n variables requires running time $2^{n(1-c_k)}$, where c_k goes to 0 as k goes to infinity. Beck and Impagliazzo (2013) proved that regular resolution cannot disprove SETH; that is, there are unsatisfiable k-CNF formulas in n variables such that each regular resolution refutation has size at least $2^{n(1-c_k)}$, where c_k goes to 0 as k goes to infinity. We give a different/simpler proof of such a lower bound based on the known characterisations of width and size in resolution, and our technique indeed works for a proof system stronger than regular resolution. The problem of finding k-CNF formulas for which we can prove such strong size lower bounds in general resolution is still open.

4.3 SAT Solvers and Proof Complexity

Sam Buss (University of California – San Diego, US)

This talk is a survey about proof complexity and Satisfiability (SAT) solvers. We first cover the exponential time hypothesis (ETH) and the strong exponential time hypothesis (SETH), abstract proof systems, and the Frege and extended Frege proof systems. We then discuss different resolution proof systems including tree-like and regular, and their relationships with the SAT algorithms DPLL and CDCL as well as pool resolution and regWRTI. It concludes with a discussion of the D-RAT verification method and its relationship with extended resolution.

4.4 Compilation of CNF-formulas: Lower and Upper Bounds

Florent Capelli (University Paris-Diderot, FR)

Joint work of Simone Bova, Florent Capelli, Stefan Mengel, Friedrich Slivovsky
Main reference S. Bova, F. Capelli, S. Mengel, F. Slivovsky, "Knowledge Compilation Meets Communication Complexity", in Proc. of the 25th Int'l Joint Conf. on Art. Intelligence (IJCAI'16), pp. 1008–1014, AAAI Press, 2016.
URL http://www.ijcai.org/Abstract/16/147

In this talk, we review recent results obtained in collaboration with Simone Bova, Stefan Mengel and Friedrich Slivovsky on compilation of CNF-formulas. The aim of knowledge compilation in this case is to transform the input CNF-formula into a succinct data structure that can be queried efficiently to solve various problems such as decision, counting or enumeration.

We start by showing how we can use tools from communication complexity to prove that CNF-formulas cannot always be compiled into succinct DNNF, a family of restricted boolean circuits that will be presented in Stefan Mengel's talk. Our result does not rely on complexity hypotheses such as $\mathbf{P} \neq \mathbf{NP}$. Having established this negative result, we then explain how the structure of the formula can be used to compile it succinctly in many cases.

4.5 A Class of Hard Formulas for QBF Resolution

Leroy Chew (University of Leeds, GB)

Joint work of Leroy Chew, Olaf Beyersdorff, Mikolás Janota
Main reference O. Beyersdorff, L. Chew, M. Janota, "Proof Complexity of Resolution-based QBF Calculi", in Proc. of the 32nd Int'l Symposium on Theoretical Aspects of Computer Science (STACS'15), LIPIcs, Vol. 30, pp. 76–89, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015.
URL http://dx.doi.org/10.4230/LIPIcs.STACS.2015.76

Proof systems for quantified Boolean formulas (QBFs) provide a theoretical underpinning for the performance of important QBF solvers. However, the proof complexity of these proof systems is currently not well understood and lower bound techniques in particular are missing. We show the hardness of the prominent formulas of Kleine Büning et al. [1] for the strong expansion-based calculus IR-calc. This, along with the strategy extraction technique, allows us to show all strict separations for the known QBF resolution calculi.

References

1 Kleine Büning, H., Karpinski, M., Flögel, A.: Resolution for Quantified Boolean Formulas. Information and Computation, Vol. 117(1), pp. 12–18 (1995).

4.6 Adding Unsafe Constraints to Improve the Performance of SAT Algorithms

John Franco (University of Cincinnati, US)

For many families of SAT formulas, the difficulty in solving an instance escalates exponentially with increasing instance size. A possible reason for this is that inferred contraints that reduce search space significantly are learned too late in the search to be effective. One attempt to control this is to add safe, uninformed constraints that are obtained from an analysis of the problem or the structure of the formula: symmetry breaking constraints, for example. This approach proves effective in some but not all cases. We propose an alternative approach which is to add unsafe, uninformed constraints early on to reduce search space breadth at shallow depth and then retract those constraints when the search breadth is still small and will not get much bigger as the search continues. By 'unsafe constraint' we mean a constraint that may eliminate one or more satisfying assignments – hence there is a risk that all assignments of a satisfiable instance may be eliminated.

We show, for example that in the case of formulas for solving van der Waerden number $W(2,6)$, adding unsafe constraints produces a bound that turns out to be $W(2,6)$. Knowledge of this bound and the conjecture that it was $W(2,6)$ was eventually used by Kouril to custom

design a solver that could prove definitively the value of $W(2,6)$. Notable is the fact that the unsafe constraints are obtained from an analysis of solutions to smaller instances of the van der Waerden family and not from an analysis of the structure of the formulas or problem properties.

4.7 Minimal Distance of Propositional Models

Miki Hermann (Ecole Polytechnique – Palaiseau, FR)

Joint work of Mike Behrisch, Miki Hermann, Stefan Mengel, Gernot Salzer
Main reference M. Behrisch, M. Hermann, S. Mengel, and G. Salzer, "Minimal Distance of Propositional Models", arXiv:1502.06761v1 [cs.CC], 2015.
URL https://arxiv.org/abs/1502.06761v1

We investigate the complexity of three optimisation problems in Boolean propositional logic related to information theory: Given a conjunctive formula over a set of relations, find a satisfying assignment with minimal Hamming distance to a given assignment that satisfies the formula (Next Other Solution, NOSol) or that does not need to satisfy it (Nearest Solution, NSol). The third problem asks for two satisfying assignments with a minimal Hamming distance among all such assignments (Minimal Solution Distance, MSD).

For all three problems we give complete classifications with respect to the relations admitted in the formula. We give polynomial time algorithms for several classes of constraint languages. For all other cases we prove hardness or completeness regarding APX, polyAPX, or equivalence to well-known hard optimisation problems.

4.8 Practical Proof Sytems for SAT and QBF

Marijn J. H. Heule (University of Texas – Austin, US)

Several proof systems have been proposed to verify results produced by satisfiability (SAT) and quantified Boolean formula (QBF) solvers. However, existing proof systems are not very suitable for validation purposes: It is either hard to express the actions of solvers in those systems or the resulting proofs are expensive to validate. We present two new proof systems (one for SAT and one for QBF) which facilitate validation of results in a time similar to proof discovery time. Proofs for SAT solvers can be produced by making only minor changes to existing conflict-driven clause-learning solvers and their preprocessors. For QBF, we show that all preprocessing techniques can be easily expressed using the rules of our proof system and that the corresponding proofs can be validated efficiently.

4.9 Linear Resolution – an Update

Jan Johannsen (LMU München, DE)

Joint work of Sam Buss, Jan Johannsen

Linear Resolution is a refinement of propositional resolution that is notoriously difficult to understand. We report on the state of our knowledge about its complexity, providing some new upper bounds and some structural properties of the system. In particular, we show that it is preserved under restrictions if and only if it is equivalent to full resolution.

4.10 Look-ahead for Solving Hard SAT Problems

Oliver Kullmann (University of Swansea, GB)

Joint work of Marijn J. H. Heule, Oliver Kullmann, Victor W. Marek

The boolean Pythagorean Triples problem has been a longstanding open problem in Ramsey Theory: Can the set $N = 1, 2, \ldots$ of natural numbers be divided into two parts, such that no part contains a triple (a, b, c) with $a^2+b^2 = c^2$? A prize for the solution was offered by Ronald Graham over two decades ago. We solve this problem, proving in fact the impossibility, by using the Cube-and-Conquer paradigm, a hybrid SAT method for hard problems, employing both look-ahead and CDCL solvers. An important role is played by dedicated look-ahead heuristics, which indeed allowed to solve the problem on a cluster with 800 cores in about 2 days. Due to the general interest in this mathematical problem, our result requires a formal proof. Exploiting recent progress in unsatisfiability proofs of SAT solvers, we produced and verified a proof in the DRAT format, which is almost 200 terabytes in size. From this we extracted and made available a compressed certificate of 68 gigabytes, that allows anyone to reconstruct the DRAT proof for checking.

4.11 Partial Polymorphisms and the Time Complexity of SAT Problems

Victor Lagerqvist (TU Dresden, DE)

Joint work of Peter Jonsson, Gustav Nordh, Magnus Wahlström, Bruno Zanuttini
Main reference P. Jonsson, V. Lagerkvist, G. Nordh, and B. Zanuttini, "Strong Partial Clones and the Time Complexity of SAT Problems", J. of Computer and System Sciences, Vol. 84, pp. 52–78, 2017.
URL http://dx.doi.org/10.1016/j.jcss.2016.07.008

The generalized SAT(S) problem is the computational decision problem of determining whether a conjunctive formula over the constraint language S is satisfiable. Even though all **NP**-complete SAT(S) problems are polynomial-time interreducible, there appears to be a vast difference in their worst-case time complexity. The question that we will concentrate on is how to explain this phenomenon using the language of universal algebra. For this purpose it is possible to associate each constraint language to a set of partial functions, so-called partial polymorphisms, satisfying certain closure properties. It has been proven that the

partial polymorphisms of a constraint language S determine the complexity of SAT(S) up to $O(c^n)$ time complexity, where n denotes the number of variables in a given instance. Unfortunately, the resulting theory is highly complex, and we will look at some unavoidable theoretical limitations of this approach. Despite this, non-trivial results can be obtained. We will give a brief survey of some of these results, and then look at how partial polymorphisms can be used to obtain kernelization procedures for SAT(S). In particular we will concentrate on SAT(S) problems admitting kernels with a linear number of constraints, and see how partial polymorphisms can be used to characterize such languages.

4.12 An Overview of QBF Reasoning Techniques

Florian Lonsing (TU Wien, AT)

We give an overview of techniques to solve quantified Boolean formulas (QBFs). At the beginning of QBF solving in the late 1990s, two main solving approaches emerged: backtracking search and expansion of variables. Backtracking search is a QBF-specific variant of the DPLL algorithm for propositional logic (SAT), called QDPLL. Variable expansion relies on the successive elimination of variables from a QBF until the formula reduces to true or false. Conflict-driven clause learning (CDCL) has been successfully adapted from SAT to QBF, resulting in the QCDCL algorithm. Analogously to resolution in CDCL, Q-resolution is the theoretical foundation of QCDCL. Unlike in SAT solving, where CDCL is the dominating approach, in QBF solving QCDCL is complemented by variable expansion. Modern implementations of expansion-based QBF solvers apply the principle of counterexample guided abstraction refinement (CEGAR). Recently, it has been shown that, from a proof complexity point of view, Q-resolution and expansion are orthogonal approaches. This theoretical result confirms related experimental observations and motivates further research in QBF proof complexity and its implications on the design of QBF solvers in practice.

4.13 QBF Proof Complexity – an Overview

Meena Mahajan (The Institute of Mathematical Sciences, India, IN)

How do we prove that a false QBF is indeed false? How big a proof is needed? The special case when all quantifiers are existential is the well-studied setting of propositional proof complexity. Expectedly, universal quantifiers change the game significantly. Several proof systems have been designed in the last couple of decades to handle QBFs, starting from the most basic Q-Resolution and Expansion+$\forall$-Reduction and going up to Frege+$\forall$-Reduction. Lower-bound paradigms from propositional proof complexity cannot always be extended – in most cases feasible interpolation and consequent transfer of circuit lower bounds works, but obtaining lower bounds on size by providing lower bounds on width fails dramatically. A new paradigm with no analogue in the propositional world has emerged in the form of strategy extraction, again allowing for transfer of circuit lower bounds. This talk will provide a broad overview of some of these developments.

4.14 Resolution and the Binary Encoding of Weak Pigeonhole Principles

Barnaby Martin (Durham University, GB) and Stefan Dantchev

We study the Resolution refutations of exponentially weak Pigeonhole Principles under both the normal and binary encodings of the stipulation that each pigeon must go in some hole. We prove that the minimal size of a Resolution refutation is $2^{\Omega(n/\log n)}$ in the binary encoding, contrasting with $2^{O(\sqrt{n \log n})}$ in the normal encoding. This is remarkable, since in tree-like Resolution the binary encoding is the easier to refute.

4.15 Approaching Backdoors in Two Non-Classical Logics

Arne Meier (Leibniz Universität Hannover, DE) and Irena Schindler (Leibniz Universität Hannover, DE)

Joint work of Johannes Fichte, Arne Meier, Sebastian Ordyniak, M. S. Ramanujan, Irena Schindler
Main reference J. K. Fichte, A. Meier, and I. Schindler, "Strong Backdoors for Default Logic", in Proc. of the 19th Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT'16), LNCS, Vol. 9710, pp. 45–59, Springer, 2016.
URL http://dx.doi.org/10.1007/978-3-319-40970-2_4

In this talk, we investigate the applicability of the notion of backdoors to two non-classical logics: Reiter's propositional default logic and the global fragment of linear temporal logic. For default logic, we will see that backdoors have to incorporate the ternary character of reasoning in this logic. By a slight technical obstacle, called extended literals, we show that our provided notion is well-chosen. Then, we show parameterized complexity results for backdoor set detection and evaluation in default logic which yield upper bounds of FPT, paraNP, and paraDeltaP2. Concerning linear temporal logic, the definition of backdoors here requires the incorporation of consistency of assignments. In the next step, we will see that the parameterized complexity of backdoor set evaluation behaves rather unsatisfactorily: most fragments are intractable. However, we identify a novel tractable fragment of LTL which is expressive enough to express 'safety' properties of a reactive system. The problem of backdoor set detection stays in all investigated cases fixed-parameter tractable.

4.16 An Introduction to Knowledge Compilation

Stefan Mengel (Artois University – Lens, FR)

In this talk we will give an introduction to knowledge compilation. We will give motivations, show how conditional lower bounds are shown and present some representations used in practical knowledge compilation and the knowledge compilation map. Throughout the talk we will present open questions and current challenges in the field.

4.17 Supercritical Space-Width Trade-offs for Resolution

Jakob Nordström (KTH Royal Institute of Technology – Stockholm, SE)

Joint work of Christoph Berkholz, Jakob Nordström

We show that there are CNF formulas which can be refuted in resolution in both small space and small width, but for which any small-width resolution proof must have space exceeding by far the linear worst-case upper bound. This significantly strengthens the space-width trade-offs in [Ben-Sasson '09], and provides one more example of trade-offs in the 'supercritical' regime above worst case recently identified by [Razborov '16]. We obtain our results by using Razborov's new hardness condensation technique and combining it with the space lower bounds in [Ben-Sasson and Nordström '08]. This is joint work with Christoph Berkholz.

4.18 Exact Algorithms for Satisfiability – an Overview

Rahul Santhanam (University of Oxford, GB)

We survey recent work on exact algorithms for Satisfiability, as well as popular hardness hypotheses such as the Exponential-Time Hypothesis and its variants.

4.19 The PPSZ Algorithm: Making Hertli's Analysis Simpler and 3-SAT Faster

Dominik Scheder (Shanghai Jiao Tong University, CN)

Joint work of Dominik Scheder, John Steinberger

The currently fastest known algorithm for k-SAT is PPSZ, named after its inventors Paturi, Pudlak, Saks, and Zane. It is simple to state but challenging to analyse. Paturi et al. give an elegant analysis for Unique-k-SAT, i.e., the case where the input formula has a unique satisfying assignment. Their analysis for the general case of multiple satisfying assignments is difficult and incurs an exponential loss in running time. In a breakthrough result in 2011, Timon Hertli showed that the Unique-k-SAT bound holds in the general case, too. His proof, though ingenious, is quite difficult and technical.

In this work we achieve two goals. Firstly, we greatly simplify Hertli's analysis, also making clear why it works and why simpler approaches are most likely bound to fail. We replace Hertli's involved inductive proof by one that uses basic tools from information complexity and simple coupling arguments.

Secondly, a simple consequence of our analysis is that if you can improve the PPSZ algorithm for Unique-k-SAT, then you can improve it for general k-SAT.

Combining this with a result by Hertli from 2014, in which he gives an algorithm for Unique-3-SAT slightly beating PPSZ, we obtain an algorithm beating PPSZ for general 3-SAT, thus obtaining the so far best known worst-case bounds for 3-SAT.

4.20 A Classroom Proof of the Random Walk 3-SAT Algorithm and its Practical Extension to ProbSAT

Uwe Schöning (Universität Ulm, DE)

The random walk 3-SAT algorithm (FOCS 99) has become part of textbooks and is taught in many classrooms. The purpose of this talk is to present an easier analysis of the algorithm. It is based on the fact that $P(X \leq a \cdot n)$ is equal to $[(\frac{p}{a})^a(\frac{1-p}{1-a})^{1-a}]^n$, up to polynomial factors. Here, X is a binomially distributed random variable with parameters n and p. Now let X be $\mathrm{Bin}(n, 1/2)$, and let Y be $\mathrm{Bin}(n, 2/3)$. The random walk algorithm randomly guesses an initial assignment, and then, it performs n random walk steps by selecting a clause not being satisfied under the current assignment and flipping the value of a randomly selected literal in this clause. The success probability of this algorithm (in case of a satisfiable input formula) can be lower bounded by $P(X \leq 1/3) \cdot P(Y \leq 1/3)$ which is, by the above equality, $(3/4)^n$. The algorithm was extended to ProbSAT (with Adrian Balint) for to participate in (and win) the SAT competition. For this purpose the flip probability distribution $(1/3, 1/3, 1/3)$ regarding the selected clause $\{x, y, z\}$ had to be changed to be proportional to $(f(x), f(y), f(z))$ where the function $f(x)$ is defined in terms of make(x) and break(x). By experiments it turns out that the make-value can be completely ignored, so that, in the case of 3-SAT, $f(x) = 2.5^{-\mathrm{break}(x)}$ is a good choice.

4.21 Understanding Cutting Planes for QBF

Anil Shukla (The Institute of Mathematical Sciences, India, IN)

Joint work of Olaf Beyersdorff, Leroy Chew, Meena Mahajan, Anil Shukla

We define a new complete and sound cutting plane proof system for false quantified Boolean formulas. We analyse the proof-theoretic strength of the new system. We show that it can p-simulate QU-resolution (and therefore Q-resolution), and indeed is exponentially stronger than these systems. However, it is incomparable (under a natural circuit complexity assumption) to even the core expansion-based QBF proof systems. On the other hand, we show that it is exponentially weaker than the QBF proof system based on Frege (introduced by Beyersdorff et al. ITCS'16). We also establish two lower bound techniques for our new system: strategy extraction and feasible interpolation.

4.22 Isomorphism of Solution Graphs.

Jacobo Torán (Universität Ulm, DE) and Patrick Scharpfenecker

Main reference P. Scharpfenecker, J. Torán, "Solution Graphs of Boolean Formulas and Isomorphism", in Proc. of the 19th Int'l Conf. on Theory and Applications of Satisfiability Testing (SAT'16), LNCS, Vol. 9710, pp. 29–44, Springer, 2016.
URL http://dx.doi.org/10.1007/10.1007/978-3-319-40970-2_3

The solution graph of a Boolean formula on n variables is the subgraph of the hypercube H_n induced by the satisfying assignments of the formula. The structure of solution graphs has been the object of much research in recent years, since it is important for the performance of SAT-solving procedures based on local search. In this talk we concentrate on the complexity of the isomorphism problem of solution graphs of Boolean formulas and on how this complexity depends on the formula type. We observe that for general formulas the solution graph isomorphism problem can be solved in exponential time while in the cases of 2-CNF formulas as well as for CPSS formulas, the problem is in the counting complexity class $\mathbf{C}_{=}\mathbf{P}$, a subclass of **PSPACE**. In addition we prove that for 2-CNF as well as for CPSS formulas the solution graph isomorphism problem is hard for $\mathbf{C}_{=}\mathbf{P}$ under polynomial time many one reductions, thus matching the given upper bound.

4.23 Lifting SAT to Richer Theories: Bit-vectors, Finite Bases and Theory Combination

Christoph M. Wintersteiger (Microsoft Research UK – Cambridge, GB)

In this talk we take a look at lifting SAT solver technology up to higher levels of abstraction and complexity in the form of Satisfiability Modulo Theories (SMT) problems. After an overview of current conceptual work and abstract solver frameworks in the area, we discuss the example of a recently developed bit-vector solver based on the model-construction satisfiability calculus (mcSAT) and how it interfaces with other theories and solvers. Finally, we touch upon future work and open problems in this area.

5 Open problems

We give a brief account of the open problem session, and describe each of the four contributions in turn.

Meena Mahajan

This open problem relates to hardness measures for resolution proofs. Given a tree-like resolution proof, and an internal node u, let $f(u)$ denote the minimum, over all parents v of u, of the width of the clause at node v. The asymmetric width $\mathrm{width}(\pi)$ of a resolution proof π is the maximum $f(u)$ over all internal nodes u of π. It is shown in [1] that

$$\mathrm{width}(F \vdash \emptyset) \leq \mathrm{awidth}(F \vdash \emptyset) + \max\{\mathrm{awidth}(F \vdash \emptyset), \mathrm{width}(F)\} - 1\,,$$

shaving +1 off the upper bound given by [2]. It remains open whether the following upper bound holds:

$$\text{width}(F \vdash \emptyset) \leq \text{awidth}(F \vdash \emptyset) + \text{width}(F) - 1\,.$$

The relation was originally conjectured in [3].

References

1 Krebs, A., Mahajan, M., Shukla, A.: Relating Two Width Measures for Resolution Proofs. Electronic Colloquium on Computational Complexity (ECCC) (2016).
2 Beyersdorff, O., Kullmann, O.: Unified Characterisations of Resolution Hardness Measures. International Conference on Theory and Applications of Satisfiability Testing (SAT), pp. 170–187. Springer (2014).
3 Beyersdorff, O., Kullmann, O.: Hardness Measures and Resolution Lower Bounds. Computing Research Repository (CoRR) (2014).

Nicola Galesi

Cutting Planes (CP) is a refutational calculus for propositional CNF formulas. The space complexity of a proof, roughly speaking, can be viewed as the amount of memory required to produce the proof.

A *memory configuration* M is a set of linear inqualities. A CP proof of I from F is a sequence $M_0, \ldots, M_k$ of memory configurations, satisfying (1) M_0 is empty, (2) $I \in M_k$, and (3) M_{i+1} is obtained from M_i by an axiom download, by inference, or by erasure. The *inequality space* of a CP refutation Π is the maximum size of memory configuration in Π.

It was shown in [1] that every unsatisfiable CNF has a CP refutation with inequality space ≤ 5, but the proof uses coefficients of exponential size. This leads naturally to the following open problem: Can every unsatisfiable CNF be refuted in CP in constant inequality space, if the coeffecients are polynomially bounded?

The next open problem concerns locality lemmas. The Locality Lemma for resolution, whose proof is trivial, states that, for any partial assignment α satisfying F, there exists a partial assignment $\alpha' \subseteq \alpha$ satisfying F such that $|\alpha'|$ is less than the space of F. A version of the Locality Lemma exists for the polynomial calculus [2, 3], and can be stated as follows. Let P be a set of polynomials, and let M be a disjoint 2-CNF with $M \vDash P$. Then there exists another disjoint 2-CNF M' such that (1) $M' \subseteq M$, (2) $M' \vDash P$, and (3) $|M'| \leq 4 \cdot \text{Sp}(P)$. We arrive at the second open problem: If we interpret P instead as a set of configurations, can we prove a version of the Locality Lemma for CP?

References

1 Galesi, N., Pudlák, P., Thapen, N.: The Space Complexity of Cutting Planes Refutations. Conference on Computational Complexity (CCC), pp. 433–447, LIPIcs (2015).
2 Alekhnovich, M., Ben-Sasson, E., Razborov, A. A., Wigderson, A.: Space Complexity in Propositional Calculus. Symposium on Theory of Computing (STOC), pp. 358–367 (2000).
3 Bonacina, I., Galesi, N.: Pseudo-partitions, Transversality and Locality: A Combinatorial Characterisation for the Space Measure in Algebraic Proof Systems. Innovations in Theoretical Computer Science (ITCS), pp. 455–472 (2013).

Oliver Kullmann

Both open problems concern the class SED of Boolean clause sets. The deficiency $\delta(F) \in \mathbb{Z}$ of a Boolean clause set F is equal to the number of clauses minus the number of variables.

The first open problem asks, simply, what is the decision complexity of SED? For the second open problem, let $\deg_F(v)$ be equal to the number of clauses in F containing variable v, and let $\mathrm{nM}(k)$ be the k^{th} non-Mersenne number. It was stated that, for a Boolean clause set $F \in \mathrm{SED}$, if $\delta(F) \geq 1$ and $\deg_F(v) \geq \mathrm{nM}(\delta(F))$ for every variable v in F, then F is satisfiable. The open problem asks whether, under these circumstances, an assignment for F can be found in polynomial time.

Stefan Mengel

Let f be the function that maps an arbitrary collection Φ of n propositional CNF formulas $\phi_1, \ldots, \phi_n$ to a string of bits $a_1 \cdots a_n$, such that $a_i = 1$ if ϕ_i is satisfiable, and $a_i = 0$ otherwise. Can f be computed in polynomial time with $o(n)$ calls to a SAT-solver? It was noted by several participants that this topic bears a close relationship to the computation of maximal autarkies.

6 Social Activities

6.1 Hike

Arne Meier (Leibniz Universität Hannover, DE)

On Wednesday at 13:45 p.m., twenty-one of the seminar participants enjoyed the great weather and friendly atmosphere during the hike. The group walked a circular route of 9.6 *km* in the direction of hill *Schafkopf*, crossed the stream *Prims*, passed the rise *junger Hirschkopf* on roughly our half-way point and reached after a long curve at *Buttnicher Straße* to eventually finish back at *Schloss Dagstuhl*. The net walking time was two hours and 22 minutes and we had an elevation gain of 140 *m*. However, we were not in a hurry. Including breaks we arrived back at approximately 16:00 p.m. – perfectly timed to enjoy the deserved cake!

6.2 Musical Evening

Joshua Blinkhorn (University of Leeds, UK)

On Wednesday evening, beginning at 8:00 p.m., all seminar participants were welcome to attend the musical evening, which took place in the castle's music room. Prior to the event, any and all participants with musical tendencies were invited to contribute a performance to the programme, either as a solo act, or – in the spirit of collaboration – as a group.

In total, seven musicians took to the stage in an eclectic collection of performances, presenting music from the Baroque and Classical eras, some well-known jazz standards, and a handful of popular songs and instrumental pieces. Making use of the instruments and sheet music provided at Dagtuhl, the concert hosted several solo performances, featured an instrumental duo, and was closed by a jazz quartet.

Being a well-attended event, the hour-or-so of music was well-received by the audience, with warm applause for each contribution. The evening was organized and compèred by Jan Johannsen (LMU München). The programme is reproduced below.

Johannes Schmidt (piano)	Goldberg Variation (J. S. Bach) Türkischer Marsch (Mozart)
Dominic Scheder (piano) and Ilario Bonacina (flute)	Suite for Flute and Piano (J. S. Bach) Vieilles Danses (B. Bartók)
Jacobo Torán (guitar)	Milonga (J. Buscaglia)
Florent Capelli (guitar and voice)	La Javanaise (S. Gainsbourg) Paris 42 (L. Aragon, L. Leonardi)

Joshua Blinkhorn (guitar and voice)	Kid Charlemagne (W. Becker, D. Fagen)
Joshua Blinkhorn (guitar), Florent Capelli (voice), Jan Johannsen (saxophone) and Dominik Scheder (piano)	Summertime (G. Gershwin) Watermelon Man (H. Hancock)

Participants

- Olaf Beyersdorff
University of Leeds, GB
- Joshua Blinkhorn
University of Leeds, GB
- Ilario Bonacina
KTH Royal Institute of Technology – Stockholm, SE
- Sam Buss
University of California – San Diego, US
- Florent Capelli
University Paris-Diderot, FR
- Leroy Chew
University of Leeds, GB
- Nadia Creignou
Aix-Marseille University, FR
- Arnaud Durand
University Paris-Diderot, FR
- Uwe Egly
TU Wien, AT
- Shiguang Feng
Universität Leipzig, DE
- John Franco
University of Cincinnati, US
- Nicola Galesi
Sapienza University of Rome, IT
- Anselm Haak
Leibniz Universität Hannover, DE
- Miki Hermann
Ecole Polytechnique – Palaiseau, FR
- Marijn J. H. Heule
University of Texas – Austin, US
- Edward A. Hirsch
Steklov Institute – St. Petersburg, RU
- Kazuo Iwama
Kyoto University, JP
- Jan Johannsen
LMU München, DE
- Peter Jonsson
Linköping University, SE
- Oliver Kullmann
Swansea University, GB
- Victor Lagerqvist
TU Dresden, DE
- Florian Lonsing
TU Wien, AT
- Meena Mahajan
The Institute of Mathematical Sciences, India, IN
- Barnaby Martin
Durham University, GB
- Arne Meier
Leibniz Universität Hannover, DE
- Stefan Mengel
Artois University – Lens, FR
- Jakob Nordström
KTH Royal Institute of Technology – Stockholm, SE
- Steffen Reith
Hochschule RheinMain, DE
- Rahul Santhanam
University of Oxford, GB
- Dominik Scheder
Shanghai Jiao Tong University, CN
- Irena Schindler
Leibniz Universität Hannover, DE
- Johannes Schmidt
Jönköping University, SE
- Uwe Schöning
Universität Ulm, DE
- Anil Shukla
The Institute of Mathematical Sciences, India, IN
- Sarah Sigley
University of Leeds, GB
- Stefan Szeider
TU Wien, AT
- Jacobo Torán
Universität Ulm, DE
- Heribert Vollmer
Leibniz Universität Hannover, DE
- Christoph M. Wintersteiger
Microsoft Research UK – Cambridge, GB

Report from Dagstuhl Seminar 16382

Foundations of Unsupervised Learning

Edited by

Maria-Florina Balcan[1], Shai Ben-David[2], Ruth Urner[3], and Ulrike von Luxburg[4]

1 **Carnegie Mellon University, US,** ninamf@cs.cmu.edu
2 **University of Waterloo, CA,** shai@uwaterloo.ca
3 **MPI für Intelligente Systeme – Tübingen, DE,** ruth.urner@tuebingen.mpg.de
4 **Universität Tübingen, DE,** luxburg@informatik.uni-tuebingen.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 16382 "Foundations of Unsupervised Learning". Unsupervised learning techniques are frequently used in practice of data analysis. However, there is currently little formal guidance as to how, when and to what effect to use which unsupervised learning method. The goal of the seminar was to initiate a broader and more systematic research on the foundations of unsupervised learning with the ultimate aim to provide more support to practitioners. The seminar brought together academic researchers from the fields of theoretical computer science and statistics as well as some researchers from industry.

Seminar September 18–23, 2016 – http://www.dagstuhl.de/16382
1998 ACM Subject Classification I.2.6 Learning, H.3.3 Information Search and Retrieval
Keywords and phrases Machine learning, theory of computing, unsupervised learning, representation learning
Digital Object Identifier 10.4230/DagRep.6.9.94

1 Executive summary

Ruth Urner
Shai Ben-David

The success of Machine Learning methods for prediction crucially depends on data preprocessing such as building a suitable feature representation. With the recent explosion of data availability, there is a growing tendency to "let the data speak itself". Thus, unsupervised learning is often employed as a a first step in data analysis to build a good feature representation, but also, more generally, to detect patterns and regularities independently of any specific prediction task. There is a wide rage of tasks frequently performed for these purposes such as representation learning, feature extraction, outlier detection, dimensionality reduction, manifold learning, clustering and latent variable models.

The outcome of such an unsupervised learning step has far reaching effects. The quality of a feature representation will affect the quality of a predictor learned based on this representation, a learned model of the data generating process may lead to conclusions about causal relations, a data mining method applied to a database of people may identify certain groups of individuals as "suspects" (for example of being prone to developing a specific disease or of being likely to commit certain crimes).

However, in contrast to the well-developed theory of supervised learning, currently systematic analysis of unsupervised learning tasks is scarce and our understanding of the subject is rather meager. It is therefore more than timely to put effort into developing solid foundations for unsupervised learning methods. It is important to understand and be able to analyze the validity of conclusions being drawn from them. The goal of this Dagstuhl Seminar was to foster the development of a solid and useful theoretical foundation for unsupervised machine learning tasks.

The seminar hosted academic researchers from the fields of theoretical computer science and statistics as well as some researchers from industry. Bringing together experts from a variety of backgrounds, highlighted the many facets of unsupervised learning. The seminar included a number of technical presentations and discussions about the state of the art of research on statistical and computational analysis of unsupervised learning tasks.

We have held lively discussions concerning the development of objective criteria for the evaluation of unsupervised learning tasks, such as clustering. These converged to a consensus that such universal criteria cannot exist and that there is need to incorporate specific domain expertise to develop different objectives for different intended uses of the clusterings. Consequently, there was a debate concerning ways in which theoretical research could build useful tools for practitioners to assist them in choosing suitable methods for their tasks. One promising direction for progress towards better alignment of algorithmic objectives with application needs is the development of paradigms for interactive algorithms for such unsupervised learning tasks, that is, learning algorithms that incorporate adaptive "queries" to a domain expert. The seminar included presentations and discussions of various frameworks for the development of such active algorithms as well as tools for analysis of their benefits.

We believe, the seminar was a significant step towards further collaborations between different research groups with related but different views on the topic. A very active interchange of ideas took place and participants expressed their satisfactions of having gained new insights into directions of research relevant to their own. As a group, we developed a higher level perspective of the important challenges that research of unsupervised learning is currently facing.

2 Table of Contents

3 Overview of Talks

3.1 Linear Algebraic Structure of Word Meanings

Sanjeev Arora (Princeton University, US)

Joint work of Sanjeev Arora, Yuanzhi Li, Yingyu Liang, Tengyu Ma, Andrej Risteski, Kiran Vodrahalli

What does a word – or more generally, a piece of text – mean? While a precise answer is difficult, many approaches involve a distributional view of semantics. I will give a 30-min survey of this area focusing on use of word embeddings. Our papers give theoretical explanations of why word embeddings exhibit linear algebraic structure even though they are derived from nonlinear methods. A more recent discovery of ours shows that different senses of a polysemous words reside in linear superposition inside the word embedding, which has implications for use of word embeddings in linguistics tasks as well as fMRI studies of the brain, as I'll sketch.

Based upon joint works with Yuanzhi Li, Yingyu Liang, Tengyu Ma, Andrej Risteski, Kiran Vodrahalli.

3.2 Interactive Clustering

Pranjal Awasthi (Rutgers University – New Brunswick, US)

Joint work of Pranjal Awasthi, Maria-Florina Balcan, Konstantin Voevodski
Main reference P. Awasthi, M.-F. Balcan, K. Voevodski, "Local algorithms for interactive clustering", arXiv:1312.6724 [cs.DS], 2014.
URL https://128.84.21.199/abs/1312.6724v2

Clustering is typically studied in the unsupervised learning setting. But in many applications, such as personalized recommendations, one cannot reach the optimal clustering without interacting with the end user. In this talk, I will describe a recent framework for interactive clustering with human in the loop. The algorithm can interact with the human in stages and receive limited, potentially noisy feedback to improve the clustering. I will present our preliminary results in this model and mention open questions.

3.3 Two recent clustering paradigms

Shai Ben-David (University of Waterloo, CA)

Joint work of Hassan Ashtiani, Shrinu Kushagra, Shai Ben-David

We consider two paradigms for semi supervised clustering. In the first, [1] the learner is allowed to interact with a domain expert, asking whether two given instances belong to the same cluster or not. We study the query and computational complexity of clustering in this framework. We consider a setting where the expert conforms to a center-based clustering with a notion of margin. We show that there is a trade off between computational complexity

and query complexity; We prove that for the case of k-means clustering (i.e., when the expert conforms to a solution of k-means), having access to relatively few such queries allows efficient solutions to otherwise NP-hard problems. In the second framework, [2], we ask the domain expert to cluster a small subset of the input data and use it to learn a metric over which k-means clustering conforms with that sample clustering. We analyze the sample complexity of that paradigm.

References

1 Hassan Ashtiani, Shrinu Kushagra and Shai Ben-David. *Clustering with Same-Cluster Queries*. Proceedings of the 30th Annual Conference on Neural Information Processing Systems (NIPS'16) 2016.
2 Hassan Ashtiani and Shai Ben-David. *Representation Learning for Clustering: A Statistical Framework*. Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence (UAI) 2015.

3.4 Questions in Representation Learning

Olivier Bousquet (Google Switzerland – Zürich, CH)

Recent successes of Deep Learning seem to rely on the ability to automatically extract and exploit structure in the data. But this process is not well understood and often ignored in theoretical analyses where the input data is treated as points in a space with some given similarity measure (which may not fully capture the internal structure of these points). However, by taking a generative point of view one can try and uncover some of the input data structure. This has led to many surprising results in image and text processing. This talk attempts to frame several recent algorithms as conditional generative density estimation and present some theoretical questions that can lead to a better understanding of representation learning.

3.5 Active Learning Beyond Label Feedback

Kamalika Chaudhuri (University of California – San Diego, US)

Joint work of Chicheng Zhang
Main reference C. Zhang, K. Chaudhuri, "Active Learning from Weak and Strong Labelers", arXiv:1510.02847v2 [cs.LG], 2015.
URL https://arxiv.org/abs/1510.02847v2

An active learner is given a hypothesis class, a large set of unlabeled examples and the ability to interactively query labels of a subset of them; the learner's goal is to learn a hypothesis in the class that fits the data well by making as few label queries as possible. While active learning can yield considerable label savings in the realizable case – when there is a perfect hypothesis in the class that fits the data – the savings are not always as substantial when labels provided by the annotator may be noisy or biased. Thus an open question is whether more complex feedback can help active learning in the presence of noise.

In this talk, I will present a feedback mechanism – when the active learner has access to a weak and a strong labeler – and talk about when it can help reduce the label complexity of active learning. If time permits, I will also discuss active learning when the annotator can say "I don't know" instead of providing an incorrect label.

3.6 A cost function for similarity-based hierarchical clustering

Sanjoy Dasgupta (University of California – San Diego, US)

Main reference S. Dasgupta, "A cost function for similarity-based hierarchical clustering", in Proc. of the 48th Annual ACM Symp. on Theory of Computing (STOC 2016), pp. 118–127, ACM, 2016.
URL https://doi.org/10.1145/2897518.2897527

The development of algorithms for hierarchical clustering has been hampered by a shortage of precise objective functions. To help address this situation, we introduce a simple cost function on hierarchies over a set of points, given pairwise similarities between those points. We show that this criterion behaves sensibly in canonical instances and that it admits a top-down construction procedure with a provably good approximation ratio.

We show, moreover, that this procedure lends itself naturally to an interactive setting in which the user is repeatedly shown snapshots of the hierarchy and makes corrections to these.

3.7 Two sample tests for large random graphs

Debarghya Ghoshdastidar (Universität Tübingen, DE)

Joint work of Debarghya Ghoshdastidar, Ulrike von Luxburg

Standard two-sample tests can achieve a high test power in the presence of large number of samples, but little is known about their performance in the small sample regime. On the other hand, it is well known that a large random graph usually concentrates about its expected (population) version. One can exploit this fact to devise two sample tests for large (inhomogeneous Erdos-Renyi) random graphs, for which a high test power can be achieved with a small population of graphs. In this talk, we will look into different variations of the problem, and present some simple tests based on matrix concentration inequalities.

3.8 Globally Optimal Training of Generalized Polynomial Neural Networks with Nonlinear Spectral Methods

Matthias Hein (Universität des Saarlandes, DE)

Joint work of Antoine Gautier, Matthias Hein, Quynh Nguyen Ngoc

We show that a particular class of non-standard feedforward neural networks can be trained globally optimal under relatively mild conditions on the data. The nonlinear spectral method

has a linear convergence rate and the conditions for global optimality can be easily checked before running the algorithm. While the algorithm can in principle be applied to neural networks of arbitrary depth, we present in the talk for simplicity the results for a one hidden layer network. The proof is based on a novel kind of Perron-Frobenius-type theorem for nonlinear eigenproblems. First experimental results show that the resulting classifiers are competitive with standard methods.

References

1 A. Gautier, Q. Nguyen Ngoc, M. Hein. *Globally Optimal Training of Generalized Polynomial Neural Networks with Nonlinear Spectral Methods*. Proceedings of the 30th Annual Conference on Neural Information Processing Systems (NIPS'16) 2016.

3.9 Multicriterion cluster validation

Christian Hennig (University College London, GB)

Main reference C. Hennig, "Clustering strategy and method selection", in Handbook of Cluster Analysis, pp. 703–730, Chapman & Hall/CRC, 2015.

Cluster validity measurement is the evaluation of the quality of a clustering, which is often used for comparing different clusterings on a dataset, stemming from different methods or with different parameters such as the number of clusters.

There are various measurements for cluster validity. Often these are used in such a way that the validity of the whole clustering is measured by a single number such as the Average Silhouette Width. But the quality of a clustering has various aspects such as within-cluster homogeneity, between-cluster separation, representation of cluster members by a centroid object, stability or within-cluster normal distribution shape, and what is most important depends on the aim of clustering. Furthermore, in many clusterings, various aspects of cluster validity differ between clusters.

In this presentation I will discuss a number of measurements of different aspects of cluster validity, partly to be evaluated for every single cluster, including some plots to summarize the measurements. A key aspect is calibration, i.e., making different measurements comparable, so that they can be used, for example, to compare different numbers of clusters. The proposed approach is to explore the variation of the index over several clusterings of the same dataset that can be generated by random clustering methods called "stupid k-means" (i.e., assigning points to a random set of centroids) or "stupid nearest neighbor" (i.e., adding nearest neighbors starting from random points).

3.10 What are the true clusters?

Christian Hennig (University College London, GB)

In much of the literature on cluster analysis there is the implicit assumption that in any situation in which cluster analysis is applied, there are some "true" clusters at which the analysis aims; and usually the "true" clustering is assumed to be unique. Benchmarking of

clustering algorithms usually is based on datasets with some assumed truth, so that it can be seen how well this truth is recovered by the algorithms.

I will argue that there are several legitimate clusterings on the same data and that defining "true" clusters is highly problematic.

I will discuss a number of related issues: philosophical background, constructive and realist aims of clustering, and various ways to define "true clusters", namely based on the data alone, on an underlying true class variable, or on probability models. Implications for cluster benchmarking and variable selection in clustering are also mentioned.

3.11 Meta-unsupervised-learning: a supervised approach to unsupervised learning

Adam Tauman Kalai (Microsoft New England R&D Center – Cambridge, US)

Joint work of Adam Tauman Kalai, Vikas Garg

Unsupervised Learning (UL) and exploratory data analysis remain one of the murkiest areas within machine learning. Theorists debate the objective of UL, and for many practical UL problems, humans dramatically outperform ML systems using prior experience in UL and prior domain knowledge or common sense acquired from prior ML tasks.

We introduce the problem of meta-unsupervised-learning from a distribution of related or unrelated learning problems. We present simple agnostic models and algorithms illustrating how the meta approach circumvents impossibility results for novel "meta" problems such as meta-clustering, meta-outlier-removal, meta-feature-selection, and meta-embedding. We also present empirical results showing how the meta approach improves over standard UL techniques for these problems of outlier removal, choosing the number of clusters and a UL neural network that learns from prior supervised classification problems drawn from the openml collection of learning problems.

3.12 Planted Gaussian Problem: Beating the Spectral Bound

Ravindran Kannan (Microsoft Research India – Bangalore, IN)

Joint work of Ravi Kannan, Santosh Vempala
Main reference R. Kannan, S. Vempala, "Chi-squared Amplification: Identifying Hidden Hubs", arXiv:1608.03643v2 [cs.LG], 2016.
URL https://arxiv.org/abs/1608.03643v2

Spectral methods can find a planted clique of size $c\sqrt{n}$ in a random graph. In spite of some effort, this is the best we know so far. Here, for a different natural problem (of a similar flavor), we show that we can do better than spectral methods.

Given an n times n matrix with i.i.d. $N(0,1)$ entries everywhere except a planted k by k submatrix which has $N(0,\sigma^2)$ entries, we show that if $\sigma^2 > 2$, then we can find a planted clique of size $o(\sqrt{n})$. We also show that if $\sigma^2 \leq 2$, no poly time Statistical algorithm can find the planted part if it is $o(\sqrt{n})$ sized. The algorithm as well as the lower bound are based on the chi-squared distance between the planted and ground densities. Some extensions will be discussed.

3.13 Recent work on clustering and mode estimation with kNN graphs

Samory Kpotufe (Princeton University, US)

Joint work of Heinrich Jiang, Samory Kpotufe
Main reference H. Jiang, S. Kpotufe, "Modal-set estimation with an application to clustering", arXiv:1606.04166v1 [stat.ML], 2016.
URL https://arxiv.org/abs/1606.04166v1

We present a first procedure that can estimate – with statistical consistency guarantees – any local-maxima of a density, under benign distributional conditions. The procedure estimates all such local maxima, or *modal-sets*, of any bounded shape or dimension, including usual point-modes. In practice, modal-sets can arise as dense low-dimensional structures in noisy data, and more generally serve to better model the rich variety of locally-high-density structures in data. The procedure is then shown to be competitive on clustering applications, and moreover is quite stable to a wide range of settings of its tuning parameter.

3.14 Proving clusterability

Marina Meila (University of Washington, US)

Joint work of Marina Meila, Yali Wan
Main reference M. Meila, Y. Wan, "Graph Clustering: Block-models and model free results", in Proc. of the 30th Annual Conf. on Neural Information Processing Systems (NIPS'16), 2016.
URL http://papers.nips.cc/paper/6140-graph-clustering-block-models-and-model-free-results
Main reference M. Meila, "The stability of a good clustering", Technical Report, 2011.
URL http://www.stat.washington.edu/research/reports/2014/tr624.pdf

Clustering graphs under the Stochastic Block Model (SBM) and extensions are well studied. Guarantees of correctness exist under the assumption that the data is sampled from a model. In this paper, we propose a framework, in which we obtain "correctness" guarantees without assuming the data comes from a model. The guarantees we obtain depend instead on the statistics of the data that can be checked. We also show that this framework ties in with the existing model-based framework, and that we can exploit results in model-based recovery, as well as strengthen the results existing in that area of research.

3.15 On Resilience in Graph Coloring and Boolean Satisfiability

Lev Reyzin (University of Illinois at Chicago, US)

Joint work of Jeremy Kun, Lev Reyzin
Main reference J. Kun, L. Reyzin, "On Coloring Resilient Graphs", arXiv:1402.4376v2 [cs.CC], 2016.
URL https://arxiv.org/abs/1402.4376v2

Inspired by notions of stability arising in the clustering literature, I will introduce a new definition of resilience for constraint satisfaction problems, with the goal of more precisely determining the boundary between NP-hardness and the existence of efficient algorithms for resilient instances. In particular, I will examine r-resiliently k-colorable graphs, which are those k-colorable graphs that remain k-colorable even after the addition of any r new edges. I will also discuss the corresponding notion of resilience for k-SAT. This notion of resilience suggests an array of open questions for graph coloring and other combinatorial problems.

3.16 Active Nearest-Neighbor Learning in Metric Spaces

Sivan Sabato (Ben Gurion University – Beer Sheva, IL)

Joint work of Aryeh Kontorovich, Sivan Sabato, Ruth Urner
Main reference A. Kontorovich, S. Sabato, R. Urner, "Active Nearest-Neighbor Learning in Metric Spaces", in Proc. of the 30th Annual Conf. on Neural Information Processing Systems (NIPS'16); pre-print available at arXiv:1605.06792v2 [cs.LG], 2016.
URL https://papers.nips.cc/paper/6100-active-nearest-neighbor-learning-in-metric-spaces
URL https://arxiv.org/abs/1605.06792v2

We propose a pool-based non-parametric active learning algorithm for general metric spaces, which outputs a nearest-neighbor classifier. We give prediction error guarantees that depend on the noisy-margin properties of the input sample, and are competitive with those obtained by previously proposed passive learners. We prove that the label complexity of the new algorithm is significantly lower than that of any passive learner with similar error guarantees. Our algorithm is based on a generalized sample compression scheme and a new label-efficient active model-selection procedure.

Sivan Sabato is supported by the Lynne and William Frankel Center for Computer Science.

3.17 Aversion k-clustering: How constraints make clustering harder

Melanie Schmidt (Universität Bonn, DE)

Joint work of Melanie Schmidt, Anupam Gupta, Guru Guruganesh
Main reference A. Gupta, G. Guruganesh, M. Schmidt, "Approximation Algorithms for Aversion k-Clustering via Local k-Median", in Proc. of the 43rd Int'l Colloquium on Automata, Languages, and Programming (ICALP 2016), LIPIcs, Vol. 55, pp. 66:1–66:13, Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, 2016.
URL http://dx.doi.org/10.4230/LIPIcs.ICALP.2016.66

There is a huge body of work on approximating clustering problems like k-median or k-means in their standard form. Less is known about the approximability of these problems once we constraint the possible solutions by, e.g., adding lower or upper bounds on the capacities of the facilities. This talk is about a side constraint that we name locality. It assumes that facilities have radii and demands that a client can only connect to a facility if it is within the facility's radius. We see how a clustering problem from game theory inspires a k-median problem with this type of constraint. This local k-median problem turns out to be surprisingly hard to approximate.

3.18 Gradient descent for sequential analysis operator learning

Karin Schnass (Universität Innsbruck, AT)

Joint work of Michael Sandbichler, Karin Schnass

We will shortly present ongoing work on analysis operator learning. We will describe the concept of co-sparsity in an analysis operator as dual concept to sparsity in a dictionary.

Based on this duality we will then propose optimization principles and associated algorithms for learning such an operator. We will show some recent results and discuss the difficulties that arise with a theoretical treatment and practical applications.

3.19 Towards an Axiomatic Approach to Hierarchical Clustering of Measures

Ingo Steinwart (Universität Stuttgart, DE)

Joint work of Philipp Thomann, Ingo Steinwart, Nico Schmid
Main reference P. Thomann, I. Steinwart, N. Schmid, "Towards an axiomatic approach to hierarchical clustering of measures", J. of Machine Learning Research, Vol. 16, pp. 1949–2002, 2015.
URL http://www.jmlr.org/papers/volume16/thomann15a/thomann15a.pdf

We propose some axioms for hierarchical clustering of probability measures and investigate their ramifications. The basic idea is to let the user stipulate the clusters for some elementary measures. This is done without the need of any notion of metric, similarity or dissimilarity. Our main results then show that for each suitable choice of user-defined clustering on elementary measures we obtain a unique notion of clustering on a large set of distributions satisfying a set of additivity and continuity axioms.

3.20 On some properties of MMD and its relation to other distances

Ilya Tolstikhin (MPI für Intelligente Systeme – Tübingen, DE)

Joint work of Carl-Johann Simon-Gabriel, Ilya Tolstikhin

Maximum Mean Discrepancy (MMD) is a metric defined on the class of probability measures and induced by a positive-definite reproducing kernel. In the recent years MMD was getting more and more attention in the ML community. In this short talk I will discuss several results on MMD, including its relation to other stronger distances like Hellinger and Total-Variation, and try to outline some of important questions for the future research.

3.21 Lifelong Learning with Weighted Majority Votes

Ruth Urner (MPI für Intelligente Systeme – Tübingen, DE)

Joint work of Anastasia Pentina, Ruth Urner
Main reference A. Pentina, R. Urner, "Lifelong Learning with Weighted Majority Votes", in Proc. of the 30th Annual Conf. on Neural Information Processing Systems (NIPS'16), 2016.
URL https://papers.nips.cc/paper/6095-lifelong-learning-with-weighted-majority-votes

Better understanding of the potential benefits of information transfer and representation learning is an important step towards the goal of building intelligent systems that are able to persist in the world and learn over time. In this talk, we discuss possible directions

for evaluating representation learning within the framework of statistical learning theory. We then focus on learning a representation from a sequence of tasks in a lifelong learning framework. We consider a setting where the learner encounters a stream of tasks but is able to retain only limited information from each encountered task, such as a learned predictor. In contrast to most previous works analyzing this scenario, we do not make any distributional assumptions on the task generating process. Instead, we formulate a complexity measure that captures the diversity of the observed tasks. We provide a lifelong learning algorithm with error guarantees for every observed task (rather than on average). We show sample complexity reductions in comparison to solving every task in isolation in terms of our task complexity measure. Further, our algorithmic framework can naturally be viewed as learning a representation from encountered tasks with a neural network.

3.22 A Modular Theory of Feature Learning

Robert C. Williamson (Australian National University)

Joint work of Daniel McNamara, Cheng Soon Ong, Robert C. Williamson
Main reference D. McNamara, C. S. Ong, R. C. Williamson, "A Modular Theory of Feature Learning", arXiv:1611.03125v1 [cs.LG], 2016.
URL https://arxiv.org/abs/1611.03125v1

Learning representations of data, and in particular learning features for a subsequent prediction task, has been a fruitful area of research delivering impressive empirical results in recent years. However, relatively little is understood about what makes a representation 'good'. We propose the idea of a risk gap induced by representation learning for a given prediction context, which measures the difference in the risk of some learner using the learned features as compared to the original inputs. We describe a set of sufficient conditions for unsupervised representation learning to provide a benefit, as measured by this risk gap. These conditions decompose the problem of when representation learning works into its constituent parts, which can be separately evaluated using an unlabeled sample, suitable domain-specific assumptions about the joint distribution, and analysis of the feature learner and subsequent supervised learner. We provide two examples of such conditions in the context of specific properties of the unlabeled distribution, namely when the data lies close to a low-dimensional manifold and when it forms clusters. We compare our approach to a recently proposed analysis of semi-supervised learning.

4 Open problems

4.1 Valid cost functions for nonlinear dimensionality reduction

Barbara Hammer (Universität Bielefeld, DE)

Nonlinear dimensionality reduction techniques have made great strides in recent years [1], and ready-to-use techniques such as the popular t-distributed stochastic neighbor embedding and efficient approximations enable a fast inspection of structure which is inherent in big

data sets [2]. These methods are not only used for interactive data inspection in striking applications e.g. from bioinformatics [3], but they have also proved valuable as a preprocessing step for high dimensional data clustering [4]. In the presentation, we will demonstrate its use for the automated contamination detection in single-cell-sequencing, an important first step in the automated analysis of data as occur in this extremely promising biotechnoloy [5]. Despite their popularity, however, means of their formal quantitative evaluation are yetr lacking. One of the probably most popular quantitative evaluation methods of nonlinear dimensionality reduction is offered by the quality framework, which quantifies the degree of neighborhood preservation of a nonlinear dimensionality reduction method in terms of a single number [6]. We will formally introduce this measure, and we will argue why it is not suited as a loss function for the evaluation of nonlinear dimensionality reduction in a learning-theoretical sense. Hence, up to our knowledge, it is open how to define a cost term for nonlinear nonparametric dimensionality reduction based on a finite set of data in such a way that it extends to a natural generalization if the number of data points is not fixed.

References

1 Andrej Gisbrecht, Barbara Hammer: Data visualization by nonlinear dimensionality reduction. Wiley Interdisc. Rew.: Data Mining and Knowledge Discovery 5(2):51–73 (2015)
2 Laurens van der Maaten: Barnes-Hut-SNE. CoRR abs/1301.3342 (2013)
3 Laczny CC, Pinel N, Vlassis N, Wilmes P. Alignment-free visualization of metagenomic data by nonlinear dimension reduction. Sci Rep. 2014; 4:4516.
4 Automated Contamination Detection in Single-Cell Sequencing, Markus Lux, Barbara Hammer, Alexander Sczyrba bioRxiv 020859; http://dx.doi.org/10.1101/020859
5 Single-cell genome sequencing: current state of the science, Charles Gawad, Winston Koch, and Stephen R. Quake, Nature Reviews Genetics 17, 175–188 (2016)
6 John Aldo Lee, Michel Verleysen: Scale-independent quality criteria for dimensionality reduction. Pattern Recognition Letters 31(14):2248–2257 (2010)

4.2 Scaling up Spectral Clustering: The Case of Sparse Data Graphs

Claire Monteleoni (George Washington University – Washington, D.C., US)

Joint work of Mahesh Mohan, Claire Monteleoni
Main reference M. Mohan, C. Monteleoni, "Effect of Uniform Sampling on Spectral Clustering", manuscript, 2016.
Main reference M. Mohan, C. Monteleoni, "A Novel Sampling Algorithm for Speeding Up the Nystrom Approximation", manuscript, 2016.
Main reference A. Choromanska, T. Jebara, H. Kim, M. Mohan, C. Monteleoni, "Fast spectral clustering via the Nystrom method", in Proc. of the 24th Int'l Conf. on Algorithmic Learning Theory (ALT 2013), LNCS, Vol. 8139, pp. 367–381, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-40935-6_26

While spectral methods for the unsupervised learning tasks of clustering and embedding have found wide success in a variety of practical applications, scaling them up to large data sets poses significant computational challenges. In particular, the storage and computation needed to handle the affinity matrix (a matrix of pairwise similarities between data points) can be prohibitive. An approach that has found promise is to instead approximate this matrix in some sense. In past work, we analyzed a variant of spectral clustering that uses the Nystrom approximation method, in which the columns are sampled uniformly. Exploiting a strong assumption of latent structure, namely that the (original) affinity matrix can be represented as block-diagonal with k blocks (or a perturbation of such), we provided bounds

on how well the clustering result approximates the result using the full-dimensional affinity matrix, with respect to the normalized-cut spectral clustering objective with k clusters.

We first pose an open question as to whether it is possible to design a sampling technique that performs better than uniform sampling, in terms of managing the tradeoff between its space and time complexity vs. the quality of the approximation. We recently provided a rejection sampling technique that addresses this goal by storing fewer and more "informative" columns. In experiments on a variety of real and synthetic data sets, our technique was able to speed up the computation and reduce the memory requirements of spectral methods, while simultaneously providing better approximations. Our observation that sparser data matrices led to decreased performance, not only for our rejection sampling technique but also for the standard uniform sampling, leads to a second open question: how to improve uniform sampling in the sparse case.

[Update: while interesting points were raised in the Dagstuhl Seminar discussion, for example, that if the matrix is sparse enough, one can avoid such sampling techniques altogether, there is still a continuum of sparsity levels which future work can address. On another note, it is worth further exploring which types of approximation guarantee on the affinity matrix imply good approximation of various spectral clustering objectives.]

Participants

- Sanjeev Arora
Princeton University, US
- Pranjal Awasthi
Rutgers University – New Brunswick, US
- Shai Ben-David
University of Waterloo, CA
- Olivier Bousquet
Google Switzerland – Zürich, CH
- Kamalika Chaudhuri
University of California – San Diego, US
- Sanjoy Dasgupta
University of California – San Diego, US
- Debarghya Ghoshdastidar
Universität Tübingen, DE
- Barbara Hammer
Universität Bielefeld, DE
- Matthias Hein
Universität des Saarlandes, DE
- Christian Hennig
University College London, GB
- Adam Tauman Kalai
Microsoft New England R&D Center – Cambridge, US
- Ravindran Kannan
Microsoft Research India – Bangalore, IN
- Samory Kpotufe
Princeton University, US
- Marina Meila
University of Washington – Seattle, US
- Claire Monteleoni
George Washington University – Washington, D.C., US
- Lev Reyzin
University of Illinois – Chicago, US
- Heiko Röglin
Universität Bonn, DE
- Sivan Sabato
Ben Gurion University – Beer Sheva, IL
- Melanie Schmidt
Universität Bonn, DE
- Karin Schnass
Universität Innsbruck, AT
- Hans Ulrich Simon
Ruhr-Universität Bochum, DE
- Christian Sohler
TU Dortmund, DE
- Ingo Steinwart
Universität Stuttgart, DE
- Ilya Tolstikhin
MPI für Intelligente Systeme – Tübingen, DE
- Ruth Urner
MPI für Intelligente Systeme – Tübingen, DE
- Ulrike von Luxburg
Universität Tübingen, DE
- Robert C. Williamson
Australian National University, AU