Program schedule

9:00 – 9:10: Welcome (Ye-Qiong Song)

9:10 – 10:30: session 1

Thomas Paris (20’ + 5’ Q&A)
Complex system modeling by composition

Florian Greff (20’ + 5’ Q&A)
Software-Defined Real-Time Mesh Networking: Protocol and Experimentation Method

Pierre-Olivier Brissaud (5’)
Intrusion/attack detection in encrypted traffic

Daishi Kondo (20’ + 5’ Q&A)
Name Anomaly Detection for ICN

10:30 – 10:45 coffee break

10:45 – 12:30: session 2

Abdulqawi Saif (20’ + 5’ Q&A)
Experimental methodology for evaluation of Big Data systems

Meihui Gao (20’ + 5’ Q&A)
Optimization models and methods for Network Functions Virtualization (NFV) architectures

Marjan Bozorg (5’)
Mixed Integer Nonlinear Programming for Membrane System Design

Xavier Marchal (20’ + 5’ Q&A)
Secure operation of virtualized Named Data Networks

Jordi Martori Adrian (20’ + 5’ Q&A)
Reliable causal delivery with probabilistic design
12:30 – 14:00 Buffet & Discussion

14:00 – 15:40: session 3

**Amina Ahmed Nacer (20’ + 5’ Q&A)**
Contribution to the secure deployment of business processes in the cloud

**Tayeb Oulad kouider (20’ + 5’ Q&A)**
Electric Vehicle Routing Problem: E-VRP

**Hoai Le Nguyen (20’ + 5’ Q&A)**
Data mining approach for studying group performance and behavior in collaborative editing

**Wazen Shbair (20’ + 5’ Q&A)**
A Privacy-Preserving Architecture for Monitoring HTTPS Services

15:40 – 16:00 coffee break

16:00 – 17:30: session 4

**Quentin Laporte-Chabasse (5’)**
Collaborative Editing Services for Peer-to-Peer Federated Systems

**Abir Ismaili Alaoui (5’)**
Improving Business Process Management through Big Data technologies and recommender systems

**Vinh Dang (20’ + 5’ Q&A)**
Trust in collaborative systems: a comprehensive study

**Victorien Elvinger (20’ + 5’ Q&A)**
Authenticated snapshot in distributed collaborative systems

**Hoang Long Nguyen (20’ + 5’ Q&A)**
End to End encrypted system for collaborative work

**Nicolas Schnepf (5’)**
Orchestration and verification of security functions for smart environments
Appendix: Abstracts of the presentations
(in order of the program schedule)

Thomas Paris, “Complex system modeling by composition”

Abstract: Complex systems are systems composed of many heterogeneous entities in interaction. For example, a smart-grid is a complex system which involves entities from at least three domains (power grid, telecommunication and information system). Such systems cannot be simply modeled as a single model, we need a multi-model: a set of several heterogeneous models (event-based, continuous...) which interact. Obviously, it is interesting to reuse preexisting checked models, but ensuring the reusability and composability of models and multi-models in a generic way is still a challenge.

MECSYCO (Multi-agent Environment for Complex System CO-simulation) is a modeling environment based on the AA4MM (Agent and Artifact For Multi-Modeling) meta-model. MECSYCO allows the description and the co-simulation of a complex system as a set of interacting heterogeneous models from different software. Regarding the reusability challenge, the upcoming step is to deal with the description of models, multi-models and experiments to improve the reusability of previous works. This implies to enhance the AA4MM meta-model with the concept of multi-models which can be reused to construct bigger multi-models (closure under coupling), and to define generic description files to store useful data for reusability. Simultaneously, these new elements allow us to deal with other complex system modeling challenges. For instance, the validity/evaluation of the multi-models or the ability to describe different decentralized and distributed simulation deployments.

Keywords: Modeling & Simulation, Composition, Complex System


Abstract: We are studying a new kind of embedded real-time network. Mesh networking of embedded components should reconcile the increasing real-time and dependability constraints, introduced by the new needs of applications in terms of bandwidth requirements and tight interactions among applications. Our approach is to allow applications to make online real-time flow resource requests and consequently dynamically allot network resources according to these requirements. To this end, additional mechanisms must be provided in order to meet the real-time constraints while the platform remains as dynamic as possible. We gather these mechanisms into a Software-Defined Real-time Network (SDRN) paradigm. I will present the principles of the SDRN protocol, which includes a credit-based traffic shaping for flow isolation, a delay and workload analysis method, a pathfinding algorithm and fault tolerance mechanism. I will also present an original hybrid network experimentation method and framework, ERICA, which we have designed to experiment on SDRN networks. It automatically generates all what is needed to conduct a specified experimentation stack in a selected environment (e.g. a simulator, testbed, ...). Hence, it allows the user to perform a high-level thinking of the whole experimentation process and reuse applications and network topologies from an experimentation stack to another one.
Pierre-Olivier Brissaud, “Intrusion/attack detection in encrypted traffic”

Abstract: The proportion of encrypted traffic increases since these last years and give users more privacy. The drawback is that it’s now impossible to use actual approaches in order to manage intrusion detection on this traffic. The state-of-the-art gives us two ways to explore in order to solve this problem: pattern detection for extracting deviant behavior and signature or use advanced proxy protocol (like man in the middle proxy) that can provide user privacy.

In the first case, because of the encryption the deep packet inspection is useless. We should handle other features that are available even with encrypted traffic. For example: packet size, interval of time between packets or packets directions, etc. With this information we can build models that can match pattern and recognize signature or deviant behavior. The other solution is to improve the work on middle boxes that can provide deep packet inspection while preserving the user privacy, keep the detection rules secret and ensure the execution of the protocol.

Daishi Kondo, “Name Anomaly Detection for ICN”

Abstract: Information-leakage is one of the main security threats in today’s Internet. As ICN is expected to become the core architecture for Future Internet, it is therefore mandatory to prevent this threat. ICN architecture relies on names and the protocol uses two different messages; Interest and Data packets. We show in this work that no information can leak through Data packets but Interest packets can be used to leak encoded crucial information with their names. As names in ICN will follow the current URL format commonly used in the Internet, we perform extensive crawling experiment of main Internet organization to obtain web URL dataset. We derive from these experiments statistic filtering technique to detect anomalous names in ICN and prevent information-leakage. This filter can be used for ICN firewall. Proposed filtering detects up to 15% of names as anomalous, i.e., statistically different from regular URL names. This work is a first step for detecting anomalous activities in ICN and preventing information-leakage.

Abdulqawi Saif, “Experimental methodology for evaluation of Big Data systems”

Abstract: Recently, Big Data phenomenon has became the hottest field in Computer Science. Big Data systems have to deal with a new class of requirements (ex: horizontal scaling, automatic failover), which obviously indicates that those systems are going to be completely different from traditional systems. This trend towards new goals motivates the market to build a lot of different- architecture solutions to generate, analyze and/or store massive data. In contrast, experimenting on Big Data systems is not yet mature due to the novelty of Big Data field. Thus, there are not enough experimenting methods for Big Data systems and this lack makes it strongly required to propose new experimenting methods.

This thesis aims at proposing new reproducible experimenting methods for (1) experimenting on Big Data storage solutions with taking into account theirs variety, and also (2) testing Big Data systems with complex infrastructure stack such as the research engine of Xilopix [1]. The first achieved contribution was about proposing a statistical
methodology to reduce the number of experiments in any experimental context while maintaining an acceptable level of results. This method is tested against Network File System (NFS) protocol in a high latency WAN environment [2]. Currently, a comprehensive study of several Big Data storage solutions is being terminated.


Meihui Cao, “Optimization models and methods for Network Functions Virtualization (NFV) architectures”

Abstract: Nowadays, telecommunication networks are populated with a large and increasing variety of proprietary hardware appliances. Therefore, it becomes more and more difficult to integrate and operate the complex hardware based appliances keeping reasonable costs. Network Functions Virtualization (NFV) is expected to overcome this problem, allowing dynamically allocation of network functions (tunneling, firewall, etc). The key problem in NFV is the service chaining problem: given a communication network in which some nodes are connected with computational servers/cloud and a set of demands asking for a sequence of Virtual Network Functions (VNFs), install VNFs on servers and route all the demands in such a way that each demand access the requested VNFs. The objective is to reduce cost/energy consumption of the system.
From an optimization point of view, the problem can be modeled as the combination of a network flow/design problem (for the demand routing part and network dimensioning) and a facility location problem (for the VNF location/server dimensioning). Both problems are widely studied in the literature, but their combination represents a new challenge.
The goal of this thesis is to propose mathematical programming models able to represent the NFV chaining problem, and to develop and compare different solution methods (exact and heuristics) to solve the problem itself.

Marjan BOZORG, “Mixed Integer Nonlinear Programming for Membrane System Design”

Abstract: Membrane networks separating multicomponent gas mixtures are becoming applicable in different aspects of industry. Membrane systems have become viable alternatives to conventional gas separation technologies (Runhong Qi, 2000).
The separation performance of a single membrane can be completely predicted by chemical engineering simulation; however, in order to increase the purity and energy efficiency, systems with more than one membrane are designed. Finding an optimal design for these systems is a complicated task, because many characteristics of the membranes must be taken into account.
The design of membrane gas separations processes can be modelled as a non-convex Mixed Integer Nonlinear Programming (MINLP) optimization problem. Continuous variables and constraints allow modeling the single membrane behavior, and discrete variables are necessary to parametrize the number and connection of different membranes (and eventually the type of membrane involved).
The goal of this thesis is to define a suitable MINLP model to cover all the aspects of membrane system configuration and to develop an ad_hoc efficient algorithm to solve the underlying optimization problem. Given the non-convexity of the problem and the large number of degrees of freedom, this task represents a challenge from both the theoretical and the computational point of view. It is a multidisciplinary project in collaboration between the chemical engineering group of LRGP and optimization team of LORIA.

Xavier Marchal, “Secure operation of virtualized Named Data Networks”

Abstract: Network operators are looking very carefully for potential opportunities and possible revenues before deploying new network equipment. This equipment is often designed for a specific usage, proprietary, and running on a specific hardware; making it very expensive to operate. Since the decision to deploy such devices follows a logic based on return of investment, this drastically limits the ambition of network operators and the innovation in their network core. For example, network operators are reluctant to globally deploy new protocols such as Named Data Networking (NDN) solutions, which proposes a novel networking paradigm where Internet data plane shifts from host-based network mechanisms to content-based ones, even if it is considered as a very promising alternative to the current IP stack.

A new trend in the networking area has recently emerged: the virtualization of network functions. NFV (Network Function Virtualization), as defined by the European Telecommunications Standards Institute, is the key technology that leverages this concept. It involves implementing network functions in software that can be run on a wide range of industry standard commodity server hardware. This initiative enables a lot of flexibility in the networks and favors innovation through the affordable, the progressive and the efficient deployment of new network functions or protocols. One of the latest research interest in the networking community is therefore to address the locks slowing down the adoption of these new standards by enabling a reliable and secure use of virtualized network equipment, which will ease the deployment of novel networking architectures and make network infrastructures more efficient. Monitoring and security are indeed primary operator requirements that need to be assured before deploying new solutions.

Jordi Martori Adrian, “Reliable causal delivery with probabilistic design”

Abstract: Ensuring reliable and ordered communication between computers usually requires acknowledgement messages. In systems with a high rate of broadcast communication, the cost of such acknowledgment messages can be large. We propose to use the causal ordering information required by some applications to detect and request missing messages. To circumscribe the number of unnecessary requests we combine local awareness and probabilistic methods. Our model and our experiments show that we obtain reliable communication within a latency equivalent to unordered communication.

Amina Ahmed Nacer, “Contribution to the secure deployment of business processes in the cloud”
**Abstract:** The constant development of technologies forces companies to be innovative in order to stay competitive. Therefore, companies are ready to outsource their business processes to the cloud to enjoy its benefits, but they are reluctant to expose their BP models which express the know-how of their companies. In fact, preserving privacy of organizations is a huge challenge. It is becoming one of the main society concerns, especially when outsourcing business processes. And probably that what is yet a problem in general, is even more exacerbated in the context of business processes (BP) because of the sensitive information they include about their organizations.

In this optic, our work consists on developing approaches allowing to securely deploy BP in the cloud while preserving the BP-Know-how and decisions strategy of companies.

---

**Tayeb Oulad kouider, “Electric Vehicle Routing Problem: E-VRP”**

**Abstract:** Electric vehicles (EV) provide a major opportunity for decarbonizing transportation activities in urban cities. However, the electric car industry is still facing two weaknesses. The first one is the limited EV driving range and the second weakness is related to the long charging time of EV. In our work, we focused on the target of managing operations of electric vehicle fleet in the context of urban logistic considering weaknesses related to electric vehicles. More precisely, our objective is developing solutions for the electric vehicle routing problem considering additional constraints related to charging and limited driving range of EV. Since the considered problem is NP-hard and in order to solve large instances of the EVRP, we develop a heuristic method based on the well-known "route first, cluster second" strategy used in the classical VRP problems. The first results allow us to confirm the contribution of partial charging in a better managing operation of electric vehicle fleet. Indeed, by allowing partial charging, we enable a larger and more realistic set of solutions for our problem.

---

**NGUYEN Hoai Le, “Data mining approach for studying group performance and behavior in collaborative editing”**

**Abstract:** Collaborative editing (CE) allows a group of people to modify a shared document simultaneously (real-time) or non-simultaneously (asynchronous). It has gained in popularity with several free-services and tools designed to support collaborative editing such as Google Drive, Wikipedia and Version control systems. Collaboration traces collected from these systems represent a valuable data for research.

Our objective is to study group performance and behavior in CE by analyzing collaboration traces. The questions we address are: How often do conflicts happen in CE? How does automatically-resolved process support CE? How do users react to conflicts during CE? Answering these questions is important to ensure good performance and user experience.

Each research question needs to be considered in both real-time and asynchronous context. Compared to asynchronous CE, in real-time CE, each collaborator receives immediately updates of other users and is therefore aware of the parts of the document being edited. However, when the number of collaborators increases, delays
for seeing other users modifications are larger and conflicts still happen. Moreover, in real-time collaboration, performances are considered more important than in asynchronous collaboration.

So far we analyzed the different effects of conflicts and user behaviors in asynchronous CE on collaboration traces that used Git.

Wazen M. SHBAIR, “A Privacy-Preserving Architecture for Monitoring HTTPS Services”

Abstract: HTTPS traffic is expanding rapidly alongside the need of Internet users to benefit from security and privacy. However, the continuous increase of HTTPS undermines the effectiveness of standard monitoring approaches. Existing solutions have privacy issues related to the decryption of the traffic in the middle. Thus, our research question is to find efficient HTTPS monitoring techniques that don’t decrypt HTTPS traffic. Firstly, we investigated the latest technique for HTTPS traffic monitoring that is based on the Server Name Indication (SNI) field of TLS. We showed that this method has weaknesses that make it unreliable for security monitoring. Thus, we improve this method by verifying the claimed value of SNI by relying on a trusted DNS server. Results show the ability to remediate the shortage of SNI-based monitoring.

Secondly, we proposed a framework to identify the accessed HTTPS services from a traffic dump, without relying neither on any header field nor on the payload content. The framework is built based on a new set of statistical features combined with machine learning algorithms and a multi-level classification approach. Recently, we study the real-time identification of HTTPS services. We propose improvements to a machine learning method to monitor HTTPS services in real-time. By extracting statistical features on the first few packets, we can identify HTTPS services very early. Experiments and a prototype implementation show that our method offers a good accuracy and satisfies the real-time requirement.

Quentin Laporte-Chabasse, “Collaborative Editing Services for Peer-to-Peer Federated Systems”

Abstract: In the context of collaborative editing, many applications adopt a centralized architecture where an authority ensures communications between collaborators. This architecture raises scalability and privacy issues. Indeed, in most cases user has no means to control implemented policies.

With the recent technological advances such as WebRTC, it is now possible to establish peer-to-peer connections between Web browsers. Therefore one can provide a peer-to-peer communication paradigm dedicated to distributed collaborative applications running in a Web browser. Such decentralized approach aims to solve both issues previously raised but brings new ones related to secure communications and interorganizational collaborations.

Interorganizational federation aims to ensure resources sharing between all the involved parties while complying to the federation’s policies. Common approaches for peer-to-peer communications used topologies based on random graphs and gossip protocol allowing the establishment of large-scale collaborative sessions.
The first research question I would like to address at the beginning of my thesis is to understand whether constraints caused by the federation could be compliant with properties introduced by unstructured networks.

Keywords: Peer-to-peer, Federation, WebRTC

Abir Ismaili Alaoui, “Improving Business Process Management through Big Data technologies and recommender systems”

Abstract: Faced with a competitive and a continuous changing environment, traditional approaches that treat a company as a closed environment are no longer appropriate. To overcome this problem of isolation and non-communication, organizations tend to increasingly use business process management-BPM. Recently with the rise of big data and its 4V (Volume, Variety, Velocity and Veracity), organizations are faced with many factors that generate real changes in the traditional BPM. Such data must be adequately exploited to extract high added value that can assist the organization in its decision making process. However, traditional business processes improvement methods such as business processes Reengineering (BPR) present different limits, as they do not allow real time use of knowledge extracted from this data by business processes.

By studying several business process models we realize that business processes are in general considered as blind, stateless, and reactive, which mean that in each business process execution we do not take into consideration neither the results from last process instances nor the context. Several researches have been done in this area to link Big data to BPM by using for example process mining or ubiquitous computing. In our research work we propose another technique based on big data analytics and recommender systems to enhance and to monitor business process models, in order to develop state full, context-aware, predictive and proactive business process models.

Vinh DANG, “Trust in collaborative systems: a comprehensive study”

Abstract: We conduct research about trust in large-scale collaborative systems where multiple users from different physical locations target common goals and interact on shared tasks. Trust between users is essential for the success of these systems. We study trust from three perspectives: objective trust, i.e. trust uniquely based on user activities in the system without including any other users opinions; subjective trust, i.e. trust based on what users think about other users; and influence of trust score, i.e. how can trust score benefit users in collaboration contexts. For studying objective trust, we focus our work on a popular platform, i.e. Wikipedia, and a behavioral studying tool, i.e. trust game. We proposed two different methods to measure quality of Wikipedia articles. We presented a novel trust computational method for repeated trust games. For studying subjective trust, we focus on signed directed social networks and proposed new methods to predict trust/distrust relations between users. All our proposed methods are validated against real-world and popular public datasets. For studying influence of trust, we organized experiments with real users.
We showed through experimentation that our method is better that the state of the art. We also confirmed the boosting effect of trust score availability in collaborative activities. Future plans include deeper study, wrapping up and writing thesis.

Victorien Elvinger, “Authenticated snapshot in distributed collaborative systems”

Abstract: In distributed collaborative systems, where a large number of participants collaborate together in a peer-to-peer manner, participants modify their own copy of the shared data. During collaboration, those copies continuously diverge then converge depending on the exchange of the participants' contributions. An adversary can conceal divergence of the copies by issuing malicious contributions. Thus, enforcing convergence of the honest participants' copies is a key requirement [4, 1].

Authenticated log-based protocols [3, 5] make misbehavior of malicious participants tamper-evident and might ensure convergence of copies owned by honest ones [5]. While they suit well the real-time settings of peer-to-peer collaboration, they generate high overheads when sub-groups work independently before merging their contributions. Snapshot-based and patch-based systems are more efficient in this last setting. However, their secured variants [6, 2] have large trust assumptions that limit both liveness and safety of collaboration [5].

We propose a protocol that combines both patch-based and log-based approaches. It enables authenticating a patch of contributions using a pruned log. Each honest participant embeds a fingerprint of the state of their copies within their messages. Two participants that merge their distinct sub-group contributions exchange a patch and the related pruned log. Each one is then able to verify the authenticity of the received patch thanks to the fingerprints they trust.

References

Long Nguyen, “End to End encrypted system for collaborative work”

Abstract: In the scope of my thesis, we focus on designing an end-to-end encrypted (E2EE) system for collaborative work among users from different enterprises that
maintain their own security servers. The system must protect communication over network from eavesdroppers, including ISPs, system providers and even the enterprises themselves while being scalable and easy to use.

The challenges are: (1) establishing an autonomous, scalable key managem-ent system which prevents adversaries to modify users public keys or provide different keys to different users and (2) constructing a multi-party key negotiation scheme that is suitable for group communication facing changes in group membership and key revocations.

In the first-year of my PhD we addressed the first challenge by designing a public log scheme, a tamper-evident structure that can be publicly auditable, to ensure the trustworthiness of enterprise key servers. The scheme efficiently and continuously monitors key servers to detect spurious keys or equivocation actions from the servers. Unlike other works which rely on complicate, expensive gossiping mechanisms, our scheme publishes log root values to a blockchain, thus reducing overhead and improving system reliability. Lastly, we introduced a web-of-trust model on top of the key management scheme to allow users effectively distinguish legitimate entities from rogue ones by taking into account opinions from other peers as well as the entities' past actions.

Nicolas Schnepf, “Orchestration and verification of security functions for smart environments”

Abstract: The development of the Internet has led to the large-scale deployment of mobile and smart environments, which are exposed to multiple security attacks. In particular, their applications coming from official or non-official markets may serve as an attack vector and have access to sensible data about users. While security solutions have been proposed to protect these environments, their deployment is often expensive with respect to scarce resources of these devices in terms of energy, cpu and memory. Cloud infrastructures and the programmability of their resources offer new perspectives for providing security functions. This PhD thesis aims at proposing, developing and evaluating new methods and techniques for building, configuring, and verifying service chains for supporting the security of smart environments. A first axis of the work will concern the design of an orchestrator capable of chaining security functions for these environments. A second axis will more specifically focus on the verification of service chains that are generated by the orchestrator with respect to high-level security properties. Once the service chaining is verified, a third axis will focus on the efficient deployment and execution of these chains in the infrastructure.