PH.D. SCHOOL IN INFORMATION ENGINEERING - UNIVERSITY OF PADOVA

FINAL REPORT

Student name: **Anna Valeria Guglielmi**

Cycle: **XXX**

Curriculum: *ICT*

Supervisor name: **Leonardo Badia**

Thesis title (final): Network science for IoT

PART 1 - COURSES, CONFERENCES AND MOBILITY

Courses for Ph.D. students

- The FFT and its use in digital signal processing (5 CFU), Prof. Silvano Pupolin
- Real-time systems and applications (5 CFU), Prof. Gabriele Manduchi
- Statistical methods (6 CFU), Prof. Lorenzo Finesso
- Applied machine learning in biomedicine (4 CFU), Prof. Enrico Grisan
- Bayesian machine learning (5 CFU), Prof. Giorgio Maria Di Nunzio

Summer schools, short courses, tutorials, seminars

- *Shaping 5G*, Federico Boccardi, 26 Marzo 2015, seminario, DEI.
- *When multimedia meets control: use of soft-real time techniques for control design*, Luigi Palopoli, 24 Aprile 2015, seminario, DEI.
- *Mathematical Scientific Challenges of 5G*, Merouanne Debbah, 3 Giugno 2015, Distinguished Lecture, DEI.
- *Coordinated Multi-Point Schemes for Interference Management in 5G networks*, Paolo Baracca, 8 Giugno 2015, seminario, DEI.
- *Do brains compute?*, Rodolphe Sepulchre, 18 Giugno 2015, Distinguished Lecture, DEI.
- *An Information-Theoretic Framework for Opportunistic Social Networks*, Tamer El Batt, 10 Settembre 2015, seminario, DEI.
- *Bitcoin, an attempt at a separation of money and state*, Pratap Pattnaik, 25 Settembre 2015, Distinguished Lecture, DEI.
- Summer School of Information Engineering, Luglio 2015, Bressanone.
- *Arithmetic for rooted trees,* November 6, 2015, Fabrizio Luccio.
- *Lovelace Test: Verso macchine creative,* December 10, 2015, Francesca A. Lisi.
- *Multifunctional organic field-effect transistors as a technological platform for sensing and biodiagnostics,* December 10, 2015, Stefano Toffanin.
- *Formal verification of complex control systems,* January 18, 2016, Alessandro Abate.
- *Neuroscience day @DEI,* February 2, 2016.
- *The Integration of Computational Medical Modeling and 3D Rapid Prototyping with Pre-Surgical Planning,* April 13, 2016, Paolo Gargiulo.
- *Low Power Wide Area Networks for the IoT,* May 4, 2016, Mark Kelly.
- *Single and Multiobjective clustering,* May 18, 2016, Ujjwal Maulik.
- *Tactile Internet,* May 18, 2016, Toktam Mahmoodi.

LeB

- *Cross layer Sensing, Estimation & Control in Wireless Networks,* June 10, 2016, Nicolò Michelusi.
- *Networks: brain, health, and society,* September 29, 2016, Maurizio Corbetta.
- Weekly internal SIGNET meetings.
- Lipari Summer School on Computational Complex and Social Sciences, July 10-17, 2016.

Participation to International Conferences and Workshops

- IEEE CAMAD, Settembre 2015, Guildford, UK.
- Italian Networking Workshop, San Candido, January 13-16, 2016.
- WoWMoM conference, June 21-24, 2016.

Other learning activities

- Tutor for the bachelor course "Telecomunicazioni" and for the master course "Elaborazione numerica dei segnali".

Mobility periods

- 15/03/2017-28/04/2017, BIOTEC Biotechnology Center TU Dresden, Germany, under the supervision of Dr. Carlo Vittorio Cannistraci, "Study of brain functional connectivity by means of already existing toolbox (EEGLAB, eLORETA) using EEG data".
- 15/05/2017-10/10/2017, Donald Bren School of Information and Computer Sciences, Irvine, California, United States, under the supervision of Dr. Marco Levorato, "Game theoretic approach to Mobile Computing"

PART 2 - RESEARCH ACTIVITY

The advent of the Internet in the 70s can be seen as the first building block for the spread and development of telecommunication networks. Afterwards, in the 80s and 90s further advancements were made thanks to the revolution in personal communications brought by mobile phone networks. Nowadays, communication networks pervade every Information Communication Technology (ICT) system: the Internet, mobile networks, smart grids, and smart transportation systems are just some examples. In the next future, we can foresee further applications involving, for examples, network virtualization, distributed assessment, and the transfer of energy via radio.
Before the emergence of the Internet and other data networks, telecommunication had a clear meaning: the telephone (and before it the telegraph) was an application of technology that allowed people to communicate at a distance by voice (and earlier by encoded electronic signals), and telephone service was provided by the Public Switched Telephone Network (PSTN). Nowadays, with the term telecommunication we refer, in general, to the transmission of signals, messages, sounds and images or intelligence of any nature by wired, radio, optical, or other electromagnetic systems. Indeed, more recently, communications at a distance have expanded to include e-mail, instant messaging, Web browsing, video conferencing, data transport, and several forms of distributed collaboration, allowed by transmission media that have further expanded from traditional copper wires to include microwave, terrestrial wireless, hybrid fiber/coaxial cable, satellite, and broadband fiber transport.

The general exchange of information between communication participants, either electrically over physical media, such as cables, or via electromagnetic radiation, involves the use of a technology, and, usually, we

LeB

refer to telecommunications, i.e., in its plural form, because it entails many different technologies. In this context, a telecommunications network can be defined as a collection of transmitters, receivers, and communcations channels that send or are used to send general messages to one another.
The main goal of the telecommunications is to provide reliable, affordable, fair, widespread devices and technologies that allow people all around the world to communicate, work together, solve problems, or just keep in touch. Consequently, several mathematical and engineering methods and models have been properly designed to address standard issues related to telecommunications aim.

One of the most addressed problems, which we may deal with when considering a transmission of information content, is errors detection and errors correction. It usually holds that information contents are divided into several packets in order to allow to the different nodes in the network to share a communication channel in a non-deterministic manner (i.e., packet swithching). In particular, Automatic Repeat Request (ARQ) is an error control technique exploiting feedback information about faulty reception of some packets to trigger their retransmissions. The most efficient basic implementation of the ARQ mechanism is the Selective Repeat ARQ (SR-ARQ) that specifically retransmitts only those packets who are reported to be in error. Recently, there has been a renewed interest for ARQ, also including hybrid ARQ techniques, where plain retransmission schemes are coupled with Forward Error Correction (FEC). ARQ-based error control is used in high-quality multimedia applications and included in the evolution of the Universal Mobile Telecommunications System (UMTS) towards High Speed Packet Access (HSPA) and the so-called recent Long Term Evolution (LTE).

It is meaningful to investigate on multimedia contents, particularly video, in order to understand their behavior and analytically characterize their performance since they are expected to be the dominant traffic component over future generation networks.
Unfortunately, most of the researches on multimedia transmission just resort to simulation, without strong analytical support for systematic evaluations. Regarding this, we believe that the reason for multimedia content being difficult to characterize in an analytical manner is related to the incremental nature of source coding that is applied to such flows. Indeed, video flows are normally obtained through incremental encoders such as those defined by Moving Picture Experts Group (MPEG). As a result, it is worth noting that different packets have different roles within the video flow and, therefore, the performance of an error correction scheme, such as ARQ, cannot be simply described as a residual error probability. From a networking standpoint, it matters whether the missing packets belong to an independently encoded, rather than an incremental frame.

It is further unclear how to apply ARQ-like techniques to such flows. A counter-argument that is often raised against retransmission-based error control for video is that it would cause delay to grow, which is not acceptable for real-time content, such as video content.
In such a background, our objective is to propose a mathematical model exploiting discrete-time Markov chains to represent the SR-ARQ retransmission mechanism for multimedia content. This aspect requires to keep into account interdependences of packets in the source coding, and to further evaluate their impact at the receiver's side in case a packet is missing, which will also be differently evaluated depending on the kind of packet. One of the main hurdles for ARQ evaluations is that classical investigations do not directly apply as the flow consists of different types of packets. Nevertheless, by means of the use of a synthetic description of the inherent structure of the packet group (i.e., the alternation of independently encoded packets and dependently encoded packets), that we call macroscopic representation, we are able to greatly simplify the system analysis. Furthermore, another side purpose is to disprove the claim according to which retransmission-based error control would cause a growing delay for video. Indeed, we show that ARQ can be applied quite easily if selective retransmissions are carefully chosen.

The telecommunications world is not only limited to ensuring the correct transmission and reception of any kind of information contents by means of a communication channel. Indeed, another recent telecommunications aim is to advance and promote significant technology progresses

eB

in green communications and networks including wireline, optical, and wireless communications and networks. In this setup, green communications and networking means sustainable, energy-efficient, energy-aware, and environmentally aware communications and networking. Therefore, innovations, new technologies, new concepts, and new principles toward a sustainable ICT are needed. The major topics of interest includes (but are not limited to) green wireline, optical, and wireless communications and networks; network and physical layer design, strategies, algorithms, protocols, and scheduling that consider environmental factors; energy-efficient and energy-aware heterogeneous networks, self-organized, and low-power sensor networks; energy efficiency in machine-to-machine communications, cooperative communications, and smart grid networks; energy harvesting, storage, and recycling for network cross-layer optimization; environmentally-aware designs of communications and networking devices and systems; and communications and networking for environmental protection monitoring.

Among these, nowadays electric vehicles (EVs) networks are raising significant interest in the scientific research community. Indeed, recently alarming pollution levels and increasing world-wide oil demand are two major circumstances that lead several countries to search for alternate energy sources and decrease carbon dioxide emissions, especially from car transportation, which makes use almost exclusively of internal combustion vehicles. In this context, EVs represent a promising technological choice to decrease pollution and reduce fossil fuel consumption. Nevertheless, their limited autonomy, long recharging times, and scarceness of recharging points pose challenges that prevents them from being suitable for many car users. Indeed, a phenomenon named range anxiety may arise: customers are wary of practicality, and, as a matter of fact, EVs market becomes a niche market. As a consequence, public investments for a better service are discouraged, and this, in turn, further decreases the attractive aspect of EVs for the average customer. Several practical solutions are sought to avoid this problem. Above all, it may be thought of enhancing the battery capacity, reducing the charging time, or increasing the frequency of charging points. In this context, our aim is to discuss how these aspects can be integrated by a proper Markov model, thereby offering a neat analytical solution to investigate all these problems.

Another significant aspect of a network in a general telecommunications sense is that it might be made up of different entities that refer to different technologies and which aim may be to provide or require different services with likely different goals. Therefore, keeping the network intelligence spread throughout the entire network is both a low-cost management solution and the proper way to involve the increased computational power of communication devices. Consequently, the need for scalability of modern communication networks has led to the practical establishment of several distributed management algorithms.

To mathematically characterize this aspect, it might be useful to consider a joint combination of several different theories that may be not directly related to the standard set of telecommunications methods and models. Basically, one of the purposes now is related to the further evaluation of classical telecommunications issues according to a more general standpoint which considers the implementation of procedures that are far away from the standard procedures used to deal with these problems because of a different context of application, but that work towards the same goal.

As a practical example, it is well known that queueing system theory develops to predict behaviors of systems subject to randomly arising demands. This was also the spirit of the early contributoins by Erlang in 1909, the works by Pollaczek and Khinchine in the 1930s, and subsequent studies. Today, queueing theory finds many applications in management of communication networks or air traffic, planning of manufacturing systems, computer program scheduling, and facility dimensioning. However, the idea to use a mathematical theory to make predictions on the behavior of multiple agents is also shared by game theory, which is now becoming more and more commonly applied to telecommunications problems. It will be discussed in the thesis that, differently from standard queueing theory, which still adopts a system view, game theory is even more extreme in considering that individual agents, called players, interact according to their particular interests that may be different for every one of them.

For this reason, we consider the joint application of both theories to a scenario including multiple candidate servers, whose objective is to provide some services to their clients that, e.g., can be thought to be packets arriving at the system for transmission. We refer to this scenario as an uplink case. Each server represents an option available at the clients and it has its own specific success probability of service completion. In this context, we model this situation as a game, where the servers are the players and they can decide to be inactive or active; staying active has a fixed local cost for the server, while a successful task completion is beneficial for the whole system. Clearly, in a system with centralized supervision, it would be best to only keep active the best server, which means the one with the highest successful probability. However, this is hindered by two facts: (i) servers act selfishly, as in any game theory setup; (ii) servers do not even know who the best server is, since they actually do not know each other's characteristics. In addition to the uplink scenario, we further consider a downlink case, where multiple candidate clients require services to some servers. Consequently, we investigate on how a server can strategically establish to which of the clients is better to offer a service making some considerations regarding the system dynamics. Indeed, for example the service required the download of a specific video, might be shared then by the clients.

To further support the idea that theories spanning different environments can be jointly combined towards a common goal, we focus on other problems properly related to telecommunications networks applying a game theoretic point of view. One of these is the so-called jamming problem, a security issue that can be seen as another application of game theory to wireless networks. Indeed, it describes a situation where a legitimate transmitting network is contrasted by a malicious attacker whose purpose is to disrupt communications. This scenario, if properly formulated, might involves just two agents, i.e., a legitimate network user and the malicious adversary, and, consequently, can be analyzed in terms of a two-player game. We consider two different setups, such as a general wireless network and an underwater acoustic sensor network, that differ in the communication channel used by the proper nodes in the network for communicating. The purpose is to evaluate the effect of the nodes' position on the resulting equilibrium conditions and to investigate the dependence of the equilibria of the game on the actual position of the jammer itself. The game theoretic character comes into play representing the uncertainty on the position information of the nodes. These evaluations can be useful in properly determing the countermeasures (if it's possible) that might be taken against the malicious jammer in order to limit its effects on the communications.

In terms of network security, unprecedented further challenges are posed by the sudden expanding of social mobile networking. Also in this context, we might have malicious adversaries that can steal sensible information and/or disrupt services due to the connection to unprotected wireless hotspots and the interaction with other nodes in the network. Moreover, nodes are usually unaware of the purposes of the agents they are interacting with; therefore, legitimate transmitting nodes can be contrasted by malicious attackers acting, again, with the purpose of disrupting communications. For this reason, our further goal is to use game theory for the identification of malicious nodes. With respect to previous similar formulations, a wider array of action options for the network's agents is considered, notably we include a choice about whether to engage or not in packet exchanging, and also malicious activity and its prevention. As a consequence, this leads to a structured analysis of the considered game, resulting in different equilibrium conditions. More in general, regarding the application of game theory to jamming problems previously discussed, in that

LB

case we consider as the critical parameter the uncertainty of the agents' positions in the network, whereas now we focus on the likelihood that the unknown agent is malicious. Indeed, we consider a scenario in which a legitimate server being unaware of the nature of its clients has to decide whether to transmit a packet, or monitor the network, or do nothing. On the other hand, the generic client's possible set of actions is to correctly forward the packet if it's necessary and if it's not malicious, or forward it adding errors if it has a malicious nature, or, finally, just ignore the packet receiving.

In general, these evaluations based on different critical parameters used to characterize the system and the objective of the game applied, beyond the design of proper countermeasures against a malicious node, allow us to see how in some cases the malicious behavior of the nodes are tolerated since a trade-off between their presence and the effect of damages that they caused are reached.

Strongly related to what has just been discussed, but on a different matter, the problem of recognizing groups of similar nodes within a mobile social network has been considered. Indeed, as a general track it can be stated that the nodes of networks, especially social networks, most of the time can be classified according to either some attributes that they show, or preferences that they have. These attributes/preferences can be seen, e.g., in terms of friendship relationship with other nodes as we consider in our investigation. However, due to the fact that it may happen to have imperfect communication (e.g. because of limited communication range), each node has just a partial view of the network and it is aware only of its relationships with the immediate neighbors. Depending on how large the communication range actually is, the problem can be in principle solved by clustering nodes by classifying the friendship relationships. Our analysis shows how a simpler procedure based on exchange of social information in a collaborative fashion is able to achieve better results than more expensive clustering algorithms. Even better, these two approaches can be combined by applying a clustering after the exchange of a given amount of social information, thus improving the overall results. This approach can be enriched considering the presence in the network of a given number of malicious nodes, and, this is the key factor that allows us to link this analysis to the previous investigations. The intent of these malicious nodes might be, for example, to forward erroneous social information to a chosen set of nodes in the network in order to compromise the network information that this set of nodes has. With this assumption, it is still possible to evaluate the effects of transmitting/forwarding malicious information in a network.

In the final thesis, it will be further discussed and described both recent addressed problems, and guidelines for future research. These guidelines if on one hand aim to continue on the path already traced by the previous tackled research topics, on the other hand are an effort to start a new research path where a telecommunications point of view might be used to deal with problems not directly related to telecommunications networks.

Following the same path of the current research, we propose as one of the future works the investigation on Mobile Cloud Computing (MCC) and Mobile Edge Computing (MEC) by means of game theory. We will discuss in more detail at the end of the thesis MCC and MEC. However, just to give an initial general idea, it is worth saying that in such contexts the mobile device users that are in the network can decide to offload the computation of tasks to a remote cloud instead of the local execution on the device (holding for MCC), or they have the opportunity to offload the tasks computation to a server located at the edge of the network named the edge server (holding for MEC). Keeping in mind that both of them have advantages and disadvantages (e.g, the remote cloud has more computation capability with respect to the edge server, whereas the edge server is nearest to the users with respect the remote cloud), our objective is to combine MCC and MEC assuming that each mobile device user in the network has three different options for a task computation: locally on the mobile device, offload to the edge server, or offload to the remote cloud.

The analysis that will be shown is just an initial simple analysis that can be further improved if we decide to consider in our evaluations the human aspects that there are beyond the device. To do so, more sophisticated games must be addressed. For this reason, for example, it may be thought to use Evolutionary Game Theory (EGT). In general, it is the game theory application to evolving populations in biology. It defines a framework of contests, strategies, and analytics into which Darwinian competition can be modeled. Consequently, it differs from classical game theory because it focuses more on the dynamics of strategy change that is influenced by the frequency of the competing strategies in the population. Therefore, in this way the dynamics of our system can be better taken into consideration allowing us to model more accurately our system. Moreover, a further combination of game theory and reinforcement learning in terms of dynamic programming might be thought in order to find the best policies using the former and to describe the internal dynamics using the latter.

In order to apply telecommunications techniques in other domains, it is worth noting that the concept of network is not only related to an information technology context, but we find several examples in nature. Indeed, the neural system of advanced living beings, especially the human brain, can be seen as a network of nodes and links, where the former are nervous cells interacting through bio-chemical reactions over the axons that work as links. A social community can be compared to a network as well: individuals are nodes that are linked according to their friendship, or work relationships, or family membership. If we consider a broader point of view, also eco-systems are networks of species. As a consequence, network studies concern several research fields, from technology to biology. Nowadays, researchers of different scientific fields focus on relevant open problems, the solution of which is to be sought in the holistic behavior of networks. Some questions must be considered: how can the interactions between malfunctioning nodes in a complex generic network result in cancer?, how and why do we see such a quick spread in some social systems and communication systems, for example involving epidemic diseases or viruses in computer?, how can some networks work even when most of the nodes are corrupted?. These problems can involve the comparison between networks of different nature that have, however, similar structure and organization.

Since the first studies on graph theory integrated with stochastic models, scientists have referred to complex networks as graphs with non trivial characteristics and with some connection schemes between the components that are neither purely regular nor completely random. The comparison between these first studies has led to the surprising discover that many networks, starting from the World Wide Web to the metabolic system in the cell or the collaboration degree between actors in Hollywood, are dominated by a small number of nodes, called hubs, that are connected to many other areas of the same network. In this way, scale-free networks and the small world property have been proposed. In the last few years, researchers study and describe the bio-inspired networking, where architectural solutions based on the evolution of living being networks are used to find solutions suitable for complex networks even related to research fields different from biology.

In the light of this scientific evolution, we suggest a twofold objective: on one hand, to thoroughly investigate models for these particular networks to understand their structure and evaluate their performance in terms of communication networks; on the other hand, to consider the application of techniques typically related to communication networks in order to optimize complex systems, with particular reference to neural systems. This joint study may be useful to yield scientific contributions in both fields, in an interdisciplinary fashion.

WB

In general, in the literature there are several examples of analysis of the cells network structure in the living beings in terms of complex network. The interaction of these cells, in addition to their behaviors, is also investigated. Some other works focused on the study of the several malfunctioning that can affect the neural network. Indeed, in the last years there has been an exponential increase in human neuroscientific research. This is due to the extensive availability of non-invasive techniques for measuring brain activity and structure, such as neuroimaging and neurophysiological recordings, producing large datasets of spatio-temporal data. Moreover, it is generally accepted that the brain can likely be modeled as a complex network and, in the literature, there can be found several description of the relationship between the brain system and a complex network.

Recently, one of the areas that is raising significant interest is that of brain connectivity; one of the goals of connectivity analysis is to make the most of the rich datasets that neuroscientists collect by evaluating the spatio-temporal dynamics present in the data. In particular, we will focus on the investigation of functional and effective connectivity in the brain, that is on the temporal correlation in the activity of two brain regions, regardless of whether they have direct anatomic links (functional connectivity) or not, and the directed causal influences one brain area produces in another (effective connectivity). The common denominator to the different types of brain connectivity is the fact that, in all instances, the system can be characterized and described as a network and in this way brain properties can be examined under formal mathematical network theories. As a final consideration, complex network analysis has emerged as an important tool to characterize brain connectivity and, since a description that consider every single components of a system is not suitable for complex systems, an holistic standpoint must be used in studying these aspects of the brain.

PART 3 - PUBLICATIONS

List of publications on international journals

J1. Anna V. Guglielmi and Leonardo Badia, ''Bayesian game analysis of uplink/downlink systems with multiple candidate agents," under preparation, 2017.

J2. Anna V. Guglielmi and Leonardo Badia, ''Exchange of social information to improve group recognition in mobile networks," under preparation, 2017.

List of publications on conference proceedings

C1. Anna V. Guglielmi, and Leonardo Badia, "A Markov Analysis of Automatic Repeat Request for Video Traffic Transmission" IEEE WoWMoM, 2014.

C2. Maria Scalabrin, Valentina Vadori, Anna V. Guglielmi, and Leonardo Badia, "Jamming in Underwater Sensor Networks as a Bayesian Zero-Sum Game with Position Uncertainty" IEEE Globecom, 2015.

C3. Valentina Vadori, Maria Scalabrin, Anna V. Guglielmi, and Leonardo Badia, "A Zero-Sum Jamming Game with Incomplete Position Information in Wireless scenarios" IEEE European Wireless, 2015.

C4. Anna V. Guglielmi, and Leonardo Badia, "Bayesian Game Analysis of a Queueing System with Multiple Candidate Servers" IEEE CAMAD, 2015.

ucB

C5. Beatrice De Lio, Anna V. Guglielmi, and Leonardo Badia, "Markov Models for Electric Vehicles: the Role of Battery Parameters and Charging Point Frequency" IEEE CAMAD, 2015.

C6. Anna V. Guglielmi, and Leonardo Badia, "Social Communication to Improve Group Recognition in Mobile Networks" IEEE Globecom, 2015.

C7. Valentina Vadori, Anna V. Guglielmi, and Leonardo Badia, "Markov Analysis of Video Transmission based on Differential Encoded HARQ" IEEE WoWMoM, 2016.

C8. Anna V. Guglielmi, and Leonardo Badia, "Analysis of Strategic Security through Game Theory for Mobile Social Networks" IEEE CAMAD, 2017.

15/09/2017

Student signature

Supervisor signature