Thesis @ ANT Lab

Matteo Cesana,
matteo.cesana@polimi.it

Alessandro E. Redondi
alessandroenrico.redondi@polimi.it

December 13, 2016

Contents

1 General Information

All master theses carried out at ANTLab are on the research topics that are currently active in the lab. Staff members are directly involved in the research activities that are often supported by national or international research projects.

Full time presence at the lab is usually not required although it is strongly encouraged when the thesis work requires a continuous interaction with other members of the research team, or if the student would benefit of a tutorship for acquiring new skills on software tools or specific research methodologies.

A regular thesis (tesi) usually requires 8-10 months of full time work, while a short thesis (tesina) requires 5-6 months. These are just estimated durations and, of course, the work is considered done only when all the scheduled tasks have been completed and the results obtained are satisfactory. However, longer durations are quite unusual. A regular thesis work allows you to have up to 7 additional points wrt your Grade Point Average (normalized in 110 scale). A short thesis allows you to have up to 4 additional points wrt your Grade Point Average (normalized in 110 scale). Example: your Grade Point Average is 26.3/30; normalized in 110 scale your GPA is equivalent to 96.43/110; if your regular thesis is perfect, then your graduation mark can be as high as 101/110; if your short thesis is perfect, then your graduation mark can be as high as 100/110 (fractional grades are rounded up to the closest integer).

The difference between a regular thesis (tesi) and a short one (tesina) is not only the duration and the amount of work, but also the objectives. A regular thesis is expected to provide innovative contributions to the state-of-the-art and usually the results of the work can hardly be predicted at the beginning. A short thesis is usually a project that can also have some limited innovative contributions but within a well defined scope that has been thoroughly investigated before.

When selecting the thesis proposal you should keep in mind that the work duration and the effort required are not negligible, therefore it is strongly recommended to choose a topic that you really enjoy working on; based on your attitudes, you have to evaluate the attractiveness of the application area, and the research methodology required for the job.

Most of our previous students enjoyed working in the lab on our thesis proposals and almost all our theses have been so far very well evaluated by the graduation committees. You can ask directly to our previous students through the google group or facebook group. We usually use these groups to stay in contact with our students and help them finding a job after graduation. We also encourage you to register to the student group on networking where we post announcements on thesis proposals, stages, job offers, etc.

1.1 Research methodologies

Even if the topic and the application area are the first elements that you consider in thesis proposals, research methodology is what youll have to deal with every

day during your thesis work. As far as the research methodology is concerned, ANTLab thesis can be divided into three families:

- Protocol design and performance evaluation Theses of this type aim at designing new communication protocols or resource management strategies for specific network scenarios (like e.g. sensor networks, wireless networks, cognitive networks, etc.) and at evaluating their performance. The performance evaluation is usually based on simulation and, if possible, on analytical models. Different simulation tools are usually adopted depending on the application scenario (like e.g. ns-2, omnet++, matlab, or ad hoc tools) but in any case some software development (usually in C/C++) activity is required.

- Prototype design and implementation This second type of theses has often the same goal of the previous one: designing and testing new protocols and algorithms. However, the validation of the proposed schemes and their performance evaluation are partially or totally based on their prototype implementation. Of course, this makes a big difference since practical skills are required and an attitude to work with network programming, distributed applications, operative systems is recommended. In the lab, prototypes are usually implemented using embedded boards or PCs with different operative systems, like Linux and TinyOS.

- Models and algorithms The last type of theses includes those using mathematical modeling approaches for the design, dimensioning or analysis of a wide range networking technologies. The mathematical instruments most often adopted are optimization theory (mathematical programming, heuristics, approximation schemes, bounds) and game theory (non-cooperative games, cooperative games, Nash equilibria). Even if mathematical skills are here the main requirement, some software skills may also be required for algorithm implementation and testing.

1.2 Must-Know List for Master Students

(AKA: Hints to Be the Perfect Master Student)

1. Research (and master thesis) is about to propose (pursue) innovative ideas based on what has been proposed beforehand.
    - Master Students are supposed to become the highest experts in their thesis topic, which means that, in the end, you have to teach your advisor.
    - Master Students have to keep literature under control almost daily.
    - Master Students MUST keep checking the literature throughout the whole thesis period.
2. Master thesis is about learning how to do things while doing them.

- Master Thesis is NOT about passively executing advisors commands.
- Master Students MUST have initiative and propose ideas to the advisor.
- Master thesis is not (always) only about running simulations on a protocol (algorithm) someone else is proposing.

Note that the best way of searching the literature is to use Google Scholar (scholar.google.com) and all the digital libraries of the publishers

- IEEE: http://ieeexplore.ieee.org/
- ACM: http://dl.acm.org/
- Scopus: http://www.scopus.com/
- Elsevier: http://www.sciencedirect.com/
- Wiley: http://onlinelibrary.wiley.com/

It is always better to start from books, when available, and survey/tutorial papers like those published on technical magazines (like IEEE Communications Magazine, IEEE Network, IEEE Wireless Communications, ACM Computer Communication Review, IEEE Tutorials and Surveys, etc.). Among research papers on the specific topic of your work, it is preferable to read first highly cited papers usually published in top quality journals/conferences and then the others. Don't be afraid of asking your advisor about top journal and conferences. This is a critical issue since in the latest period bogus conferences and journals are flourishing continuously. Using the most appropriate keywords for your search is of paramount importance to focus on the papers that are relevant for your work. You'll learn to identify these keywords while you'll get more and more familiar with your topic.

1.3 How to Write a Thesis

The following are just short notes on how the final writing work should be organized. Your advisor will give you more detailed indications before you start writing.

- The thesis is organized into chapters, sections and subsections exactly like a book (sub-subsections should be avoided if possible)
- Text formatting is not very important however it is preferable to use a font not too small (11pt-12pt), reasonable margins (2 cm at least), and a quite large line spacing (approx 1.5). The best thing to do is to take a look at a previous thesis. **ANTLAB strongly encourages the use of *LateX* for thesis writing and production.** Ask you advisor about LateX thesis templates which can help you in writing.

- Thesis length is not very important however very well written short thesis are preferable to long inaccurate thesis.
- Thesis can be written in English or Italian.
- Before starting the writing work you should have clearly in mind how you want to organize your thesis. Write down an index first and ask for the comments of your advisor.
- The first chapter of the thesis is always entitled Introduction and is basically an extended abstract (2-5 pages) of the whole thesis. It must be easy to ready also to non-experts and should give a rather complete picture of 1) the general topic considered, 2) the specific problem addressed, 3) the approach/solution proposed, 4) the methodology adopted, 5) the results obtained, 6) the organization of the thesis (In Chapter 2,).
- A few chapters after the introduction (usually 1-3) should be devoted to background information (general overview of the topic, overview of the relevant technologies, etc.) and literature review.
- The core chapters of the thesis are those describing the solution you propose and the results you obtained. In those chapters you must include all the information required to give the reader an accurate description of your work. If possible the chapter(s) related to the presentation of the solution proposed should be kept separated from those devoted to the numerical results.
- The last chapter is always entitled Conclusion and it is a recap of the thesis where basically the main results are summarized and commented. At the end of this chapter possible future developments are presented.
- References come right after the last chapter. They must be numbered and provide a complete information of the book/paper/document/website cited. Each reference must be cited at least once in the text. It is preferable to number references according to the first citation in the text. The formatting style of the references should be similar to that adopted in technical papers.

2 Thesis Proposals

In the following section, we report some of the available thesis proposal which can be carried out under the supervision of ANT Lab staff members. We will do our best to keep this document up-to-date, however, we warmly encourage those who are interested in thesis proposal to check directly with ANT Lab staff members for "fresh" proposals.

2.1 Data Analytics in Wireless Networks

The role of data analytics is recognized as central nowadays in many application fields. Think of the huge amount of money that Over The Top (OTT) players like Google and Facebook earn out of the data they get from their customers (targeted advertisement, etc.).

Being able to collect data, process it and extract from data some type of "knowledge" is fundamental also to design and optimize communication networks. Let me convince you with a practical example. If a cellular network can leverage data on the current position of a user it is serving, then it can configure/re-configure the network elements (and/or the radio resources) to serve the user in the best possible way, e.g., by steering directional antenna beam right to the exact position of the user itself.

All the following thesis proposal have a common streamline: get some networking data, apply some type of machine learning algorithms to data, extract some knowledge, optimize the network according to such knowledge. The specific proposals differ on the specific data (reference technology), the specific algorithm used to process the data (supervised/unsupervised learning, classification, regression, clustering, etc.) and, obviously, on the type of knowledge and optimization.

2.1.1 Wi-Fi / LTE traffic prediction for mobile offloading

Mobile data offloading, often known as WiFi offloading, is the use of complementary network technologies for delivering data originally targeted for cellular networks. Offloading reduces the amount of data being carried on the cellular bands, freeing bandwidth for other users. It is also used in situations where local cell reception may be poor, allowing the user to connect via wired services with better connectivity.

In particular, Wi-Fi offloading is an emerging business domain with multiple companies entering to the market with proprietary solutions.

In this thesis, the candidate will leverage realistic data traces coming from a Wi-Fi network and a LTE cellular network located in the same place to analyse the feasibility of Wi-Fi offloading. This will include making predictions for the future load on the two networks, finding spatio/temporal correlation in the data traces and other analysis in order to optimise the offloading process. Through the analysis of network data traces, the objective of this thesis is to understand to which extent prediction is possible, and if there are substantial differences in

the load of the two co-located networks. All these findings may help in designing future hybrid Wi-Fi / LTE networks.

Suggested read-ups: https://www.ruckuswireless.com/carriers/3g-offload.

2.2 Performance Evaluation and Design of Internet of Things systems

2.2.1 Performance Evaluation of Application layer/middleware solutions for the IoT

The IoT field is characterized by diverse application scenarios with heterogeneous requirements. Such diversity is reflected in the availability of very many application layer/middleware solutions which are used to transfer information across sensor devices. In this context, the IoT designer willing to set up a new IoT product or to deploy an IoT network has often to face the dilemma on which protocol (or protocol family) fits best the reference application environment.

To name a few available protocols at the application layer, we have the COnstrained Application Protocol , the Message Queuing Telemetry Transfer (MQTT) , the XMPP etc. All these protocols differ for the specific data transfer paradigm they are based on (client/server, publish/subscribe, etc.), the specific QoS features they support and the specific lower layer protocols they are leveraging to transfer information.

The goal of this thesis is to provide a comprehensive performance evaluation of a subset of these protocols in different IoT use case. The thesis will require first to settle reference network scenarios (network topology, type of traffic, etc.) to run the performance evaluation; then, the candidate will use commercial simulation tools (e.g., ContiKi Cooja) to evaluate the performance of the aforementioned protocols in the reference scenarios. In the end, the performance evaluation will (might) be extended also to a small-scale testbed to be set up at hand.

Suggested read-ups: COAP [2], MQTT-S [7], XMPP [6].

2.2.2 Cellular IoT Performance Evaluation

Applications and services for Smart Cities often require to collect and process data coming from field sensor nodes in the urban environment. As far as the communication is concerned, data must travel from devices which are immersed in the urban environment towards information sinks, and viceversa. Generally speaking, there are three most commonly used ways to realize such communication patterns: (i) through classical mobile Cellular Networks, (ii) through IoT-Dedicated Cellular Networks, (iii) through Multi-Tier Networks. In the case of Cellular Mobile Networks, the reference architecture is the one of legacy mobile radio networks (2G/3G/4G) with a Radio Access Network (RAN) in the front end and a Core Network (CN) at the backhand. Whilst cellular mobile networks are designed to serve primarily human-to-human and human-to-machine traffic, IoT-dedicated cellular networks are taking pace to

fill in the need of designing low-cost, low-energy M2M applications with limited traffic requirements. IoT-dedicated cellular operators often share the same proposition value which includes reduced energy consumption and Total Cost of Ownership (TCO) with respect to classical cellular operators, global reach and plug-and-play connectivity. Standards implementing the IoT-dedicated cellular paradigm include SigFox, LoraWAN, Weightless and Ingenu.

The goal of this thesis is to develop a simulation framework to evaluate the performance of LoraWAN in terms of geographical coverage and serviced traffic under different operation conditions. The simulator will represent multiple cells of LoraWAN network covering several end devices in urban and rural environments.

Suggested read-ups: [1], [8].

2.2.3 Virtual Sensor Networks

Usually, WSNs are designed and deployed in a vertical, application-specific way, in which the hardware and network resources are customized to the specific application requirements. On one hand, such design paradigm allows to have optimal performance on the specific application, but, on the other hand, it precludes resources (hardware and software) reuse when other applications and services must be contemplated. In the end, this has led in the past to the proliferation of redundant WSNs deployments.

In this context, virtualization is a promising technique to target more efficient resource utilization, lower cost and increased flexibility and manageability in WSN deployments. The key idea behind virtualization is to decouple the physical infrastructure and resources from application ownerships. Thus, network virtualization technologies abstract away physical resources including node processing/storage capabilities, available communication bandwidth and routing protocols, which can then be composed at a logical level to support usage by multiple independent users and even by multiple concurrent applications.

In this thesis, we focus on the design of a virtualization engine for WSNs. Namely, we will consider a general purpose WSN which can be used to support multiple applications and we will propose a mathematical programming framework to optimally allocate virtual sensor networks to the requesting applications. The proposed framework will have to allocate the physical resources of the general purpose WSN to multiple concurrent application while accounting for the limitations in the physical resources (processing, storage, available bandwidth, limited communication range) and the specific application requirements.

The students working on this thesis should have some basic knowledge on Operations Research and optimization tools.

Suggested read-ups: [5].

2.3 Spectrum and Infrastructure Sharing in 5G Networks

The exponential growth of mobile data and the increasing diffusion of bandwidth-eager user applications is pushing the migration to more spectrally efficient mo-

bile technologies such as LTE-A and eventually 5G. However, individual network roll-outs represent large sunk costs for Mobile Network Operators (MNOs) and particularly for new entrants especially given the currently unaligned growth of data demand from MNO revenues. Sharing agreements for greenfield network deployments are an attractive alternative to cut down on the upfront infrastructure cost but also, when spectrum sharing is allowed, to benefit from aggregating spectrum resources, where the latter is essential to boost network capacity.

The goal of this thesis is to propose game-theoretic models to assess the dynamics involved in spectrum and resource sharing. A toy example: suppose that two mobile network operators VodaPhon and Felecom already have an up and running 3G network with macro cells, etc. They both have their own market share (number of customers) and they have to face the strategic decision on wether to improve their network by deploying LTE in the hope to improve their revenues.

The students should have some background on game theory and/or operations research.

Suggested read-ups: [4].

2.4 Visual Sensor Networks

Visual sensor networks (VSN) are composed of several wireless cameras able to acquire, process and transmit images and video from the surrounding environment and will play a major role in the upcoming Internet of Things paradigm. A non-exhaustive list of applications enabled by this technology includes: object / face recognition, event detection, pedestrian tracking, parking and traffic monitoring, etc...

Like traditional Wireless Sensor Networks, VSN are battery-operated and energy constrained: therefore, all operations involving VSNs must be carefully designed, optimised and implemented. In particular, tools from the Computer Vision and Image Processing areas are generally exploited in order to make such networks more efficient.

A candidate willing to write a thesis in this area should have a minimal background in signal processing (better if a course on Image / Multimedia processing was taken) and good programming skills (Matlab and/or C++).

2.4.1 Visual descriptor crossmatching

Several image processing tasks (object/face recognition, tracking, etc.) are carried out relying on visual descriptors extracted from an image. Such descriptors summarise the content of the image in a concise way, capturing only the most salient information and discarding the rest. As an example, descriptors extracted from an image with a person over a blue skye and a green field will just capture the presence and the details of the person and ignore the background.

Several descriptor algorithms are available: SIFT (Scale Invariant Feature Transform) is probably the most common and has been used in many commercial applications. However, running SIFT on a sensor node is very costly from the

energy point of view. Cheaper alternatives are available (BRISK, FREAK, etc.), at the cost of reduced performance.

This thesis will address the problem of translating a set of low-cost descriptors (e.g. BRISK) that can be extracted from a node in a set of more complex descriptors (e.g. SIFT) that can be used on a server to perform image analysis. The tools used for such translation will be taken from the machine learning area, including solution based on the novel deep learning paradigm.

Suggested read-ups: [3]

2.4.2 Lightweight car signature based on visual descriptors

Our street and highways are disseminated by smart cameras constantly monitoring car traffic. Often, it is important to recognise and identify multiple occurrences of the same cars at different times. This problem is known as 'vehicle re-identification'. Although the license plate can be used as identifier of a vehicle, in many occasions it is difficult or even impossible to use it.

In this thesis, the candidate will implement and evaluate a system for car re-identification meant for visual sensor networks. The system should be able to re-identify cars by extracting simple yet accurate descriptors from realistic traffic images, in different light conditions (night, day). At the same time, the algorithms used should be simple enough to be run on low-cost hardware, such as the one used for VSN scenarios.

Suggested read-ups: [9]

References

[1] Martin Bor, Utz Roedig, Thiemo Voigt, and Juan Alonso. *Do LoRa low-power wide-area networks scale?*, pages 59–67. ACM Press, 11 2016.

[2] Carsten Bormann, Angelo P. Castellani, and Zach Shelby. Coap: An application protocol for billions of tiny internet nodes. *IEEE Internet Computing*, 16(2):62–67, 03 2012.

[3] Antonio Canclini, Matteo Cesana, Alessandro Redondi, Marco Tagliasacchi, João Ascenso, and R Cilla. Evaluation of low-complexity visual feature detectors and descriptors. In *Digital Signal Processing (DSP), 2013 18th International Conference on*, pages 1–7. IEEE, 2013.

[4] L. Cano, A. Capone, G. Carello, and M. Cesana. Evaluating the performance of infrastructure sharing in mobile radio networks. In *2015 IEEE International Conference on Communications (ICC)*, pages 3222–3227, June 2015.

[5] Carmen Delgado, Jos Ramn Gllego, Mara Canales, Jorge Ortn, Sonda Bousnina, and Matteo Cesana. On optimal resource allocation in virtual sensor networks. *Ad Hoc Networks*, pages –, 2016.

[6] A. Hornsby, P. Belimpasakis, and I. Defee. Xmpp-based wireless sensor network and its integration into the extended home environment. In *2009 IEEE 13th International Symposium on Consumer Electronics*, pages 794–797, May 2009.

[7] U. Hunkeler, H. L. Truong, and A. Stanford-Clark. Mqtt-s: A publish/subscribe protocol for wireless sensor networks. In *Communication Systems Software and Middleware and Workshops, 2008. COMSWARE 2008. 3rd International Conference on*, pages 791–798, Jan 2008.

[8] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid. Internet of things in the 5g era: Enablers, architecture, and business models. *IEEE Journal on Selected Areas in Communications*, 34(3):510–527, March 2016.

[9] Linjie Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang. A large-scale car dataset for fine-grained categorization and verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3973–3981, 2015.