EPS
UdG Escola Politècnica Superior

Treball final de grau

Estudi: Grau en Enginyeria Informàtica

Títol: Anàlisi estadístic i visualització de la robustesa de xarxes

Document: Resum

Alumne: Jordi Capdevila Mas

Director/tutor: Josep Lluís Marzo Lázaro
Departament: Arquitectura i Tecnologia de Computadors
Àrea: Arquitectura i Tecnologia de Computadors

Convocatòria (mes/any): setembre/2015

Introduction

It is well known that we live now in a society more intercommunicated than ever. The clearest example that comes to mind is Internet, which is the huge network that connects many devices and people around the world and is changing and growing rapidly every day.
But many other types of services are required to ensure the modern lifestyle: electricity, running water, gas, public transports...

All these services are distributed through their own network, each with its characteristics, and they have become essential in our everyday life. Unfortunately, all these networks have to deal against failures in their elements everyday, which could endanger the correct service provision: a cut on a power line could provoke a blackout, a blockage on the train tracks could lead to several delays.
Due to the omnipresence of these networks, a failure could involve a big problem upon a portion of the population, that is why it is necessary to have robust networks against these failures, capable to fight against them and minimize their impact.

The study of the network robustness tries to find methods to measure this robustness, thus, being able to design better networks or to enhance the existing ones.

Motivations

Nowadays, most of the presented projects used to be about software implementation, mainly centered in web or mobile applications. This project offers the possibility to go in depth into a research area, it is therefore a kind of project different to most of the others.

Moreover, the area of network robustness is focusing several lines of research in the last decades and it has a growing importance nowadays. Most of the documentation encountered and listed in the bibliography has been done since year 2005.

Finally, being able to learn and expand the work of a PhD thesis has also been a great motivation.

Starting point and objectives

As it has been said in the last section, **this project starts from a recent PhD thesis** made by a former student of the Universitat de Girona, Marc Manzano, in collaboration with the research group BCDS[1] (Broadband Communications and Distributed Systems). Its main objective is to find a way to measure the network robustness using several metrics summed all together. In order to refer to this project more easily, from this point on it will be named as *ALPHA project*.

Starting from this, the objectives of the current project are:

- Replicate the functionalities, methodologies and concepts introduced in *ALPHA project*.

[1] `http://bcds.udg.edu`

- Once achieved the latter, thoroughly analyse its functioning and enhance and correct all the necessary aspects.
- Automatize and generalize the whole process to make it applicable to any kind of network, attack, metric...
- Implement a method to visualize the implied data and networks in the process.

ALPHA project

ALPHA project is the result of the PhD thesis of a former student of the Universitat de Girona, Marc Manzano, in collaboration with the research group BCDS. As it has been already said, **this project has served as a starting point** of the current project you have in your hands.
ALPHA project is extensively explained in the main document. It basically addresses two main goals:

- The first one is to **find a general framework to deal with all the existing metrics** related with robustness and try to find a mechanism which allows to work with several of them together. The proposed solution in this project is named *R*-value*.
- The second one is to **find a way to graphically visualize the network robustness** variability against the increase of failures.
  The proposed solution in this project is named *robustness surface*.

System requirements

According to the project objectives, these are the functional requirements needed to accomplish them, and the non-functional requirements which have to be taken into account:

Functional requirements

- **Read topologies:** be able to read topologies, saved as graph structures in a file, for its further manipulation.
- **Generate basic topologies:** generate parametrized, basic topology structures (bus, tree, grid...) to be able to experiment with them.
- **Calculate metrics:** implement needed functions to be able to compute all the required metrics on a graph.
- **Generate attacks:** be able to generate different types of attack on network elements, either nodes or edges, and affect either a percentage or a given number of graph elements.
- **Visualize topologies:** find a way to graphically represent topologies with the proper layout and also visualize the attacked elements to see the topology affectation.
- **Compute *robustness surface*:** be able to reproduce the process to calculate the *robustness surface* on a given topology. This requirement includes the previous ones.

It also includes the process to calculate PCA on a data set and the function to calculate the *R*-value*.

- **Choose the proper metrics:** find a way to choose the metrics to be calculated, according to a cost/benefit criterion.

Non-functional requirements

- **Functional and adaptable code:** make a code as functional as possible and pay attention to its quality.
  Make the whole process customizable and adaptable to different types of experiments, i.e., different types of attacks, percentage of failures, different chosen metrics, different topologies...

- **Easy interaction:** provide the system with an easy input and with a general and reusable output methods. Keep the user interaction easy but sufficient to control all the process features.

- **Efficient code:** the process involves expensive mathematical and statistical calculations, given the fact that it could work with large graphs or matrices. Give special attention to code efficiency and especially avoid unnecessary repetitive processes or minimize their cost.

Project analysis

According to the functional requirements, the implemented software environment is designed as shown in figure 1.

Functionalities are grouped into different modules and **all the user interaction is done through a general interface** module, which has the proper options to access to all the other modules functionalities. Basically, the other modules are classified in:

- **Auxiliary modules:** they basically manage the input and output data, which is specified below.

- **Functional modules:** they implement the main processes and calculations to satisfy the requirements.

The involved data is organized within the project folder structure, which is basically composed by:

- ***"resources"* folder:** it includes data related to the topologies and with the experiment declarations, defined in CSV format.

- ***"out"* folder:** it contains a subfolder for every experiment. Each subfolder contains the saved data related to the results of the main calculation processes in CSV format.

- ***"conf"* folder:** it contains a configuration file with the basic environmental and path variables. In the beginning, it is necessary to set the value of some variables of the configuration file.

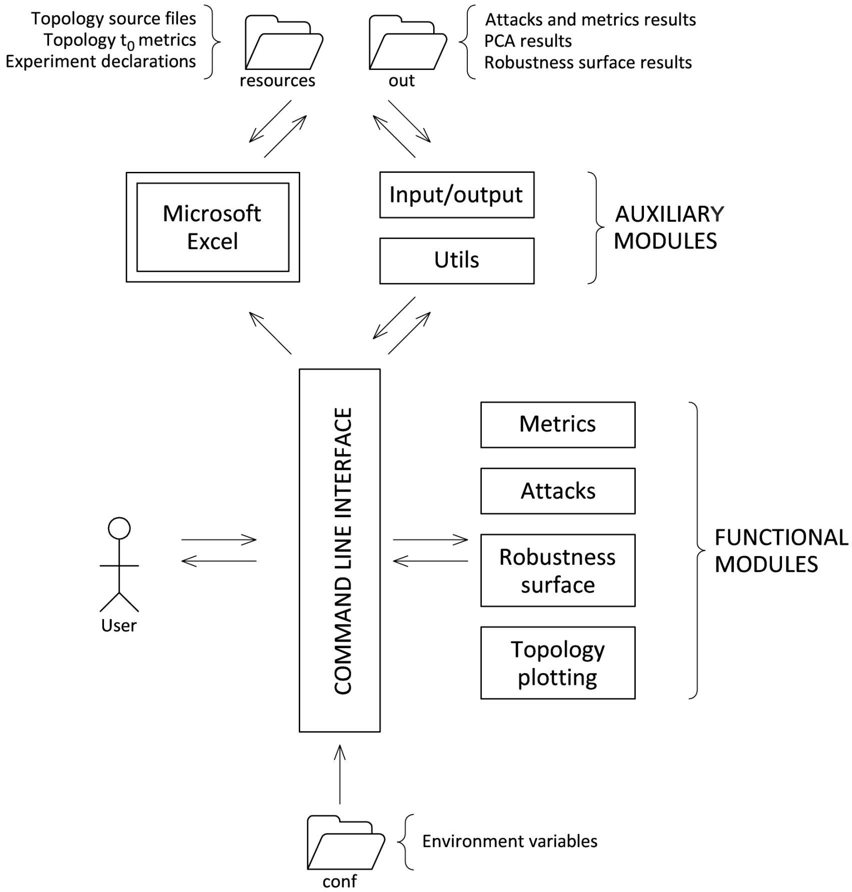

Figure 1: Project design: modular scheme analysis

Conclusions

Recalling the objectives which were set at the beginning of the project, these are my general conclusions about their achievement:

- It has been possible to **implement a software tool that emulates *ALPHA project***. The calculation processes to get *R*-value* and *robustness surface* have been fully understood and analysed, and they have been implemented with success too.
- Furthermore, the whole process has been provided with **new functionalities and methodologies**, in order to make it more generic, simple and customizable for the end-user.
- It has to be noted that **an extensive analysis work has been done** over the related concepts in robustness area, especially with all the implied metrics. A deep study of all them, their behaviour and their implication in terms of robustness has been done and all these results can be seen in the project annexes.
- The project has introduced me into a research area, providing **new concepts and knowledge.** Moreover, I have been able to apply some of the concepts which I have learnt throughout these last years, such as functional programming, software engineering patterns, statistical and graph related concepts... Even with the redaction of this document I have learnt new things, because it has been done in LaTeX.

In general, the resulting project has fulfilled the main initial expectancies and, personally, it has a satisfactory result.
However, it also has to be said that some aspects have not been achieved, or at least partially achieved, most of them due to time limitations. Maybe some of them were not main objectives, but it was also a matter of personal ambition to achieve them:

- The project has been focused first, in a theoretical research area, and then with the implementation of the latter. The next step would be to apply all of this on a practical example or a real case, in order to **translate the work to real results and obtain clearer conclusions.**

- Once the results of an experiment are obtained, the project lacks of a work to **take advantage of the potential of this data.** For the moment, the results can be *"manually"* visualized, but a set of methodologies would be appropriate to analyse this data and get conclusions to could enhance the network, for example.

- Some results comparing the performance between this project and *ALPHA project* are introduced at the end of the main document. They show that the introduced solutions in this project have changed the resulting *robustness surfaces* in some way. **A deeper analysis of the differences and the correctness of the results** would be appropriate.