Proposals for Master Degree Thesis
Research Group on
Blockchain and Social Networking
Reference: Laura Ricci
May 2017

This document proposes some topics for Master Thesis related to some research areas we are currently investigating.

- **Thesis in the area of the blockchain technology**

  1. *A blockchain-based support for social networking.*

     The recent increase in reported incidents of privacy and security breaches in social networks are mainly due the centralized nature of these neworks, which enables third-parties to collect and control massive amounts of personal data. This thesis regards the definition of a framework for the support of social networks exploiting a distributed ledger and a Distributed Hash Table (DHT). The idea is to define a protocol that turns a blockchain into an automated access-control manager that does not require trust in a third party. Unlike Bitcoin, transactions in the defined system are not financial, they are used to store, query and share social data. While content is stored on the DHT, the blockchain will register pointer to the content, together with the related access control policies. Smart contracts will be exploited to trigger spefic actions on the shared content, for instance the revoking of right the access the content in presence of some event.

     The system will be implemented by exploiting the Ethereum blockchain, which will be accessed through a local node. *This thesis investigates a research topic suitable as a basis for a PhD research proposal.*

  2. *Analysin the Bitcoin's transaction graph*

     An interesting and unique feature of Bitcoin is that the complete list of all the transactions occurred from its inception is publicly available. This enables the investigation of funds movements to uncover interesting properties of the Bitcoin economy. The purpose of this thesis is to analyze the Bitcoin users' graph for discovering if certain patterns in the graph are generated by economical phenomena. For instance, we plan to investigate if a price fluctuation of the cryptocurrency corresponds to the appearance of specific patterns in the graph. The analysis will be defined by exploiting *webgraph*, (http://webgraph.di.unimi.it/), a framework supporting the management of very large graphs, exploiting modern compression techniques.

  3. *Blockchain for IoT; in collaboration with Nokia Bell Labs, Paris*

One of the bigger trends in the blockchain world has been a need for privacy and confidentiality. This has led to all the private blockchain solutions being developed today. However, both from a global and a local perspective, this has created heterogeneous applications which have limited scopes and suffer from bootstrap issues. Indeed, thinking about the upcoming Internet of Things (IoT), having a fragmented ecosystem in which every application runs on its own infrastructure and requires its own application will not scale, for a number of reasons; the bigger one being the fact that low-end devices will not have enough power to run all the required blockchains pile. This problem has been already (partially) addressed and a solution has been proposed in the form of "merged mining". Based on this approach, unknown or poorly adopted blockchain applications can leverage the mining process of other well established solutions (such as the Bitcoin blockchain). The merged mining is a quite complex process in which transactions from auxiliary applications are ordered and their merkle trees hashed out. However, it allows users to access the whole data structure in an all or nothing approach, hence threatening peers privacy. In this thesis, we investigate the feasibility of a hypervisor for blockchains that will allow the instantiation of virtual blockchain applications—all running on the same physical infrastructure and mining algorithm. The goal is to design an hypervisor that enables peers to deploy and join auxiliary blockchain applications with an underneath global mining power, whilst being granted limited access on the hosted data structures. On the one hand virtual peers will be run by real peers (thus paving the way to additional incentive schemes) and will have a local scope while, on the other hand, they will enjoy the scalability and the security properties provided by the global blockchain infrastructure running in the background. *This thesis can be a research topic suitable as a basis for a joint PhD with Bell Labs*

- **Thesis in the Area of Online Social Networks**

  1. *Predicting users' behaviour in social networks*
     Understanding of the user temporal behaviour is a crucial aspect for all those systems that rely on user resources for the daily operations, to both reduce the maintenance cost and increase their performance. For instance, Distributed Online Social Networks exploit the devices of their users to take on and share the tasks needed to provide the service, such as storing the social data published by the users. Nowadays, the extensive use of personal devices (such as notebooks or smart phones) allow their users to explicitly connect to the DOSN in order to share information and contents. As a result, the availability of data in these systems is strongly affected (or reflected) by the temporal behaviour of their users. Linear predictors allow to predict the behaviour of users peers to foresee which of them will be online in the future. The main purpose of the project will be the investigation of adaptive techniques that allow a linear predictor to react quickly to the occurrence of specific patterns of users behaviour

or unexpected events (such as week ends or holiday days).

2. *Analysing Social Temporal Human Behaviour*

   Decentralized Online Social Networks are distributed platforms which provide the classic social services without the presence of a centralized server. Understanding the temporal behaviour of the usres is a crucial aspect for defining content availability politics in this systems. We developed a Chrome/Mozilla Firefox plugin which is able to retrieve online status of Facebook users, which decide to install the plugin, and of all their friends. In particular, we retrieve information about online status every 30 seconds. We estimate to analyse more than 3 months of users activity. This means that we will have a huge amount of information that need a postprocessing. The aim of this thesis is to analyse the logged data, to obtain insights of different aspects of the users' behaviour.

3. *Collecting and analysing Facebook Groups*

   Most of the popular Online Social Networks services available on the Internet (such as Facebook, Google Plus, etc.) allow their users to organize their friends in groups. For example, Facebook allows their users to create public or closed groups in order to facilitate contents sharing on specific topics and to link together users of the network who share particular interests (such as hobbies, school, work). Another popular OSN service is provided by Google Plus, which allow their users to organize their friends into circles, i.e., private group of contacts based on types of relationship (colleagues, family, acquaintances, etc.) or shared features (users interests or hobbies). With reference to this scenario, the candidate will extend a java-based crawler able to retrieve information about a set of Facebook groups. The main aim of the project will be the enhancement of the crawler and the analysis of the collected data through the tools for complex network analysis presented in the course.

4. *Assisting users in defining privacy policies*

   The proliferation of social networking websites open new challenges for privacy protection and studies of online social behaviour indicate that users often fail in specifying privacy policies that match their needs. A possible approach to manage this problem is to assist the users in managing their contacts according to groups (or circles) of contacts with common features (such as location, hobbies, school, etc.). The main aim of the project will be the definition of a tool to assist the users of the Online Social Networks in grouping users friends into different social circles which share similar features. For this purpose, the candidate will use a Facebook dataset which contains the profiles information of about 300.000 users.