

Proposte di Tirocini/Tesi Triennali

Laurea Triennale in Informatica

Gruppo di Ricerca: Laura Ricci

Giugno 2021

I tirocini/tesi proposti nell'ambito del nostro gruppo di ricerca riguardano principalmente a due aree di ricerca. La prima riguarda le blockchain, la seconda l'analisi di grafi derivanti da reti sociali. Vengono proposti inoltre una serie di tirocini da svolgere in collaborazione con istituzioni esterne (CNR, Università di Cambridge, CMRE La Spezia), anche in questo caso i tirocini sono supervisionati da Laura Ricci. Per ogni classe di tirocini, si riportano alcune proposte. Gli studenti interessati possono contattare Laura Ricci a laura.ricci@unipi.it per fissare una Skype call.

Ambito Blockchain

- **Analisi di strategie di gambling sulla blockchain di Bitcoin: il caso di Satoshi Dice**

Satoshi-Dice è stato un servizio di gambling attivo su Bitcoin negli anni 2010-2015. Le scommesse su tale servizio consistono nell'inviare bitcoin ad indirizzi predefiniti, dove ciascun indirizzo corrisponde ad una diversa percentuale di vincita. Esistono diverse strategie di gioco, quali martingale, D'Alambert, etc. Scopo del tirocinio è quello di individuare (tramite scraping di forum, siti web, etc.) gli indirizzi messi a disposizione da Satoshi per le scommesse, quindi analizzare la blockchain di Bitcoin per individuare pattern di transazioni relativi alle scommesse, possibilmente generati in maniera artificiale da bot.

- **Analisi di Blockchain Online Social Media**

Il successo di Steemit, piattaforma di social networking interamente implementata su blockchain, risiede in due aspetti. Un primo aspetto è quello economico, ovvero la possibilità di ricevere una ricompensa per i contenuti che si producono per la piattaforma e la possibilità di convertire la moneta ottenuta (STEEM) in Bitcoin e quindi in euro/dollari aprendo di fatto agli utenti la possibilità di ottenere un guadagno concreto per i loro contributi. Un secondo aspetto è quello sociale legato alla completa libertà di espressione, spesso adottato come cavallo di battaglia, contrapposta al continuo controllo dei servizi centralizzati come Facebook e Twitter che possono decidere di bloccare contenuti, unita

all'autogestione da parte degli utenti stessi. Queste piattaforme sono ad oggi poco conosciute e diversi loro aspetti possono essere oggetto di tirocini, che riguardano aspetti quali

- Bot detection (utilizzando anche tecniche di machine learning)
- Analisi del grafo delle transazioni, (motif e struttura)
- Valutazione dell'evoluzione della ricchezza degli utenti nel tempo
- Confronto con piattaforme analoghe (TRON, EOS, etc..).

- **Studio della piattaforma EOSIO**

EOSIO è una interessante piattaforma blockchain di nuova generazione. Dal sito eos.io, “EOSIO is the leading open-source blockchain platform that enables transparency in transactions at the speed and scale needed to solve real-world challenges. We believe transparent and decentralized systems will keep users and builders in control and help architect integrity into our world”

Questi tirocini riguardano lo studio e l'utilizzo della piattaforma EOSIO. Il primo tirocinio ha lo scopo di studiare le diverse configurazioni di blockchain disponibili in questa piattaforma (private blockchain, private multi-host blockchain, private, proxy-based blockchain) e valutarne il loro utilizzo in diversi scenari. Un secondo tirocinio ha come scopo lo studio del sistema di gestione dei permessi utilizzato dalla piattaforma EOSIO. In particolare, lo studente dovrà analizzare le caratteristiche del sistema e confrontarlo con i tradizionali sistemi di gestione dei permessi (ACL, RBAC, ABAC, LKH,..) ed, eventualmente, integrarli.

- **Sviluppo di un'applicazione che implementi il modello Self Sovereign Identity per la mobilità internazionale studentesca**

Keywords

Sviluppo - Ingegneria del software - Self Sovereign Identity - Sovrin

Contesto

Al giorno d'oggi, l'identità online è principalmente gestita da enti privati che assegnano ad una persona un'identità in un dato specifico contesto (profilo Facebook, account Steam, credenziali INPS, cliente IKEA, etc). Questo comporta ogni persona a dover gestire molteplici account ed identità, che spesso replicano le stesse informazioni (nome, cognome, data di nascita, etc), ma i dati di questa persona sono gestiti e sotto la custodia di questi enti privati (anche noti come Identity Provider).

Il modello Self Sovereign Identity (SSI) spinge verso una gestione più decentralizzata e centrata all'utente delle identità. In particolare, un utente ha assegnato una sola identità composta da vari “status” (detti claim), ognuno firmato da un Issuer. Per esempio, il comune di residenza firma i dati anagrafici, l'università i dati relativi alla professione di studente, Steam i dati relativi all'account sulla

piattaforma, etc. Quando un utente vuole usufruire di un servizio, il fornitore (Verifier) verifica che l'utente abbia un claim firmato da un Issuer di fiducia (per esempio, GameStop potrebbe accettare claim firmati da Steam, ma non da Epic Games, per rilasciare sconti in negozio).

Obiettivo

In questo progetto di tesi l'obiettivo è comprendere le potenzialità ed i limiti di un sistema SSI applicato alla mobilità internazionale studentesca, come il programma Erasmus+. Uno studente di Pisa che va in Erasmus a Londra deve necessariamente iscriversi all'università di Londra come studente. Questa azione va ripetuta per ogni destinazione nell'arco del programma Erasmus+ (nota: nel programma Erasmus+ è possibile fare più di un'esperienza all'estero per ciclo di studi).

Percorso

Lo sviluppo di questa tesi seguirà questo percorso:

1. Studio dell'architettura SSI e delle sue tecnologie, in particolare Sovrin;
2. Modellazione del caso d'uso, Erasmus+, come architettura SSI;
3. Sviluppo di un prototipo in Sovrin. In caso di problemi con Sovrin, il prototipo potrà essere sviluppato in uPort;
4. Paragone dei pregi e difetti tra il corrente metodo di iscrizione degli studenti Erasmus, ed il modello proposto nella tesi.

Riferimenti

- [Sovrin](#)
- Parte ora di [Hyperledger Indy](#)
- Un progetto simile alla proposta di tesi: [StudyBits](#)
- [Technical report](#) di StudyBits

• Analisi ad alto livello di attack graph (collaborazione con Prof. Baiardi)

Questo lavoro deve definire e validare strategie che permettano di individuare percorsi di attacco, i.e. cammini, che partano da una certa superficie e raggiungano alcuni obiettivi. Interessa sia definire la struttura dei grafi i.e. attributi associati a nodi ed archi, sia gli algoritmi da utilizzare per risolvere classici problemi di sicurezza quali valutazione del rischio e calcolo delle contromisure. E' molto importante che i grafi utilizzati permettano di operare a diversi livelli di astrazione e quindi con diverso grado di dettaglio rispetto al sistema target. Gli algoritmi sviluppati possono essere validati usando grafi prodotti da sistemi esistenti

Ambito Reti sociali

- **Multilayer social networks**

Le social network sono diventate parte integrante della vita quotidiana della maggior parte delle persone in tutto il mondo. Ne esistono di innumerevoli tipi e con i più disparati scopi: da quelle più "classiche" per rimanere in contatto con gli amici (come Facebook) a quelle orientate alla sfera lavorativa (come LinkedIn), da quelle basate sulla condivisione di immagini (come Instagram) a quelle più incentrate sulla condivisione di video (come YouTube o TikTok). Nonostante gli utenti cerchino di fare molta attenzione ai dati che condividono su ogni piattaforma, spesso collegano direttamente o indirettamente i propri account, relativi a diversi social, tra di loro. Questo ha gravi ripercussioni sulla privacy: se collego i miei account social, i dati che non condivido con una piattaforma potrebbero essere raccolti dalle altre, di fatto fornendo un completo identikit dell'utente. Scopo del tirocinio sarà quello di estendere un lavoro di tirocinio concluso nel quale sono stati collezionati dati da Facebook e Instagram al fine di costruire una rete sociale multi-livello. Questa rete sociale ha due livelli, uno per Facebook ed uno per Instagram, ed i livelli sono collegati dal fatto che un utente di una rete sociale appartiene anche all'altra. Grazie a questa rete multilivello, si cercherà di estrapolare informazioni aggiuntive per arricchire i singoli profili oppure per inferire altre corrispondenze non note.

- **COVID-19 virus and social spreading**

Obiettivo del tirocinio è quello di capire quali analogie e differenze ci sono nella diffusione delle informazioni relative al COVID-19 nelle social network e la diffusione effettiva del virus stesso tra le persone. Mentre per la diffusione effettiva del virus abbiamo a disposizione una certa conoscenza proveniente dai dati ufficiali, per la controparte relativa alle reti sociali, mancano ancora dati realmente significativi. Lo scopo del tirocinio è quello di creare un piccolo framework per la raccolta di dati su argomenti specifici (COVID-19 in questo caso), concentrandosi su contenuti che riferiscono l'argomento selezionato. Tra le varie piattaforme da considerare ci sono sicuramente Twitter [1] e Medium [2]. Non sono comunque da escludere eventuali dati già resi pubblici da alcune piattaforme [3].

Refs:

[1] <https://developer.twitter.com/en/docs>

[2] <https://github.com/Medium/medium-api-docs>

[3] <https://dataforgood.fb.com/>

- **Analisi di Gruppi FB**

Negli ultimi anni si è assistito ad un progressivo cambiamento riguardo l'utilizzo delle social network. Infatti, sempre più spesso gli account personali vengono usati per la promozione di eventi, beni o servizi, ad esempio per la promozione della propria azienda, di fatto atrofizzando l'aspetto di socialità che è stato uno dei capisaldi dei servizi di social networking. Parallelamente si è assistiti alla creazione dei cosiddetti Gruppi Sociali Online (GSO), ovvero gruppi di utenti che si uniscono in comunità legate a particolari argomenti, ad esempio, condizioni mediche o sport. Gli esempi più di successo ad oggi sono i gruppi di Facebook, le hashtag communities di Twitter e le communities di Steemit. Alcuni studi recenti stanno dimostrando come i GSO stanno diventando il principale vettore di socialità all'interno delle Online Social Network, destando l'interesse di scienziati e sociologi. Data la relativa novità è importante capire fino a che punto queste nuove funzionalità si stanno diffondendo e stanno sostituendo le vecchie modalità di socializzare online. Scopo del tirocinio è quindi quello di identificare un insieme di GSO (ad esempio quelli di Facebook) e di analizzarli in maniera estensiva, sia dal punto di vista delle interazioni degli utenti che da un punto di vista dei contenuti che circolano nel gruppo. I gruppi dovranno mostrare un'attività rilevante e per ottenere i dati sarà necessario sviluppare un crawler web di scraping utilizzando dei sistemi di automazione del browser. Dopo aver ottenuto le informazioni necessarie, saranno svolte numerose analisi per individuare le proprietà dei gruppi, concentrandosi anche sulla loro natura dinamica dei gruppi. Infine, saranno studiate informazioni testuali, come bigrammi e topic detection.

- **BOSMs in OSNs**

Negli ultimi anni abbiamo assistito ad una rivoluzione nel mondo dei social network che ha portato allo sviluppo di nuove soluzioni basate su blockchain. Queste soluzioni vengono indicate come Blockchain Online Social Media (BOSM) e hanno introdotto una importante novità, ovvero la possibilità di ricompensare gli utenti per le loro azioni (post, commenti e mi-piace). Se da un lato hanno dimostrato che possono esistere modelli differenti di social network, da un altro lato non hanno avuto, almeno per il momento, il successo necessario per imporsi sul mercato. Scopo di questo tirocinio è studiare come gli utenti delle social network centralizzate (come Facebook e Twitter) percepiscano queste nuove proposte (come Steemit). Per svolgere il tirocinio lo studente dovrà realizzare uno scraper che raccoglierà informazioni sulle social network tradizionale, che riguardano le BOSM, con lo scopo di raccogliere opinioni su queste ultime. Per raggiungere lo scopo si prevede l'utilizzo di alcune librerie di sentiment analysis disponibili online e altri tool di text analysis con l'obiettivo di capire quali sono i punti forti e quali quelli deboli delle BOSMs.

Tirocini in collaborazione con istituzioni esterne

Questi tirocini saranno svolti in collaborazione con enti esterni, ma sempre sotto la nostra supervisione (come tutor interni)

- **Blockchain solution for robot swarms - Collaborazione con CMRE (Centre for Maritime Research and Experimentation) – La Spezia.**

La tecnologia dei robot swarms sta rivoluzionando diverse applicazioni industriali. Tuttavia, diverse caratteristiche, quali la autonomia dei robot, il controllo decentralizzato, il comportamento collettivi che emerge dalle loro interazioni, rendono complessa la loro effettiva utilizzazione in sistemi industriali. Questo tirocinio indagherà come la combinazione della blockchain con altri sistemi distribuiti, come i robot swarms, può fornire gli strumenti per rendere i sistemi più sicuri, autonomi e flessibili. Il tirocinio considererà. In particolare, i robot utilizzati in ambiente marino, e questo rende l'applicazione particolarmente interessante, in quanto si tratta di progettare una blockchain in un ambiente con bassa banda di comunicazione. Il tirocinio si inquadra nell'ambito di una cooperazione tra il Dipartimento di Informatica e il CMRE di La Spezia.

- **ALPS smart licenses encoder - Collaborazione con Università di Cambridge.**

This tirocinio will be carried on in collaboration with the IIPM (Innovation and Intellectual Property Management) Laboratory of the University of Cambridge, UK, and ALPS Software Ltd. <https://www.ifm.eng.cam.ac.uk/research/innovation-and-ip-management/research-projects/smart-contract-for-licensing/>. The aim of the ALPS project is to address the unsatisfactory status quo in the largely manual and error prone Intellectual Property (IP) management field by providing an automated licensing payment and management system based on smart licensing agreements. Such goal can be achieved by employing a Distributed Ledger Technology (DLT) solution as a trusted computing platform to automatically track, secure, and enforce licensing agreements. This allows for license agreements to be cheaper, more fine grained, and more complex, as their execution is delegated to the DLT, easing the burden on the participants. Moreover, the ALPS system also allows novel business models as the license agreements visibility is extended to the entirety of the DLT. This means that, if product usage information is accessible by the DLT, then novel licensing models, e.g. outsourcing part of the royalty charges to users, become possible. At the core of the ALPS platform lies a translation

module that translates the royalty logic expressed by license agreements into a custom defined structured format. This step is needed to remove the ambiguities of a traditional human written document. Such intermediate format is then fed to a compiler that translates it into a smart contract bundle, named the smart license. To enforce formal correctness of the translation process further constraints are set on both input and output formats. Aim of this tirocinio is for the student to define a parser to automatically generate smart licensees. Familiarity with compiler and parser (either bottom up or top down) design is welcome but not necessary, as well as a basic knowledge of context free grammars (as the input of the parser can be expressed by one). The smart license can be encoded either on Solidity (to be deployed on Ethereum) or Javascript (to be deployed as chaincode on Hyperledger Fabric). The current ALPS prototype is based on a permissioned Ethereum clone, as such Solidity is the preferred output language and full support and testing capabilities will be provided mainly for it.

Requirement: The final report/thesis will need to be written in English.