P2P Systems and Blockchains

Final Projects

Academic Year 2021/2022

The student must submit*one and only one* of the following final projects.

1 TRY dApp: a dApp for the NFT Lottery

Develop a dApp which must include two components: the blockchain-based backend and the web-based frontend. The backend is implemented through the set of smart contracts of the final term, while the frontend must be implemented from scratch. The student may use web3 Javascript, Java or Python as base language.

The Dapp provides different interfaces, one for the lottery manager, and one for the lottery users. It is required to implement a function *Create Lottery* which is used by the manager to create a new lottery. This function creates an event which is notified to all the potential users who can be interested in joining the lottery. Later, another event notifies the users of the start of a new round of the lottery. It is required to implement all the functionalities needed for the lottery presented in the final term. Keep in mind that the dApp must be able not only to invoke smart contracts' functions, but also to catch events generated by the smart contracts.

Do note that, since web applications are not the main topic of the course, the user interface of the frontend of the dApp will not be taken into account for grading. As long as it is working, a basic GUI is all that is required. Still blinding colours, fixed 1024x720 windows and other examples of unusable UIs will not be appreciated.

2 Deanonymization techniques for Bitcoin

Address clustering is a technique used for breaking the pseudo anonymity of *Bitcoin*, and consists in linking addresses that are probably controlled by the same user based on the information available from the transactions registered on the blockchain. In [1] two of the most popular heuristics used to deanonymize *Bitcoin* are presented, i.e. the *Multi Input Heuristics, MIH* and the *One-time Change Heuristics, OCH*. The application of these heuristics produces a set of cluster, each one containing the *Bitcoin* addresses that probably belong to the same user.

Read carefully the paper and apply the two heuristics presented (do not consider the third heuristics proposed in the paper) to cluster the *Bitcoin*

addresses present in the transaction log which have been referred in the Mid Term.

Note that the analysis can be implemented by using any programming language or data management tools that you desire. Your solution will not be graded on efficiency, you can use, for instance, an interpreted language. However the report must clearly describe the high-level code of the algorithm used to perform the clustering.

To evaluate the effectiveness of the proposed heuristic-based address clustering algorithms, consider the following metrics introduced in [1], i.e. the ratio R of address reduction obtained by the application of the heuristics, defined as follows

$$R = \frac{|A|-|C|}{|A|}$$

where A is the set of all addresses in the transactions and C is the set of all clusters after clustering.

Evaluate the output of the heuristics considering the following metrics:

- the distribution of the size of the address clusters
- the address reduction in three scenarios, in the first one only MIH is applied, in the second one only OCH, in the final both heuristics are applied.
- the temporal trend of the address reduction over different time scales. Consider the address reduction at regular intervals, for instance after K blocks, with K stepping on an interval of values.

Is this clustering method accurate? List at least one potential source of false positives (clustering addresses which are not actually owned by the same entry) and one source of false negatives (failing to cluster addresses which actually are owned by the same entry) in this method. What strategies could you use to make your clustering more accurate?

3 Project Submission Rules

The project must be developed individually. The material to be submitted for the evaluation is the following one:

- a report (pdf document) describing the main features of the project. The report should include: a brief summary of the project choices and implementation, and, for the second project, a set of plots evaluating the heuristics.
- a pdf document reporting the code of the deveoped applications.

The report and the code must be submitted electronically, through the Moodle. The project will be discussed a week after its submission. The discussion of the project consists in the presentation of a short demo, which can be run on the personal laptop and a general discussion of the choices made in the implementation of the system.

I recall that the oral examination is waived for the the students who have passed the Mid and Final Term. Do not hesitate to contact us for any doubt (laura.ricci@unipi.it).

References

[1] G. S. Manku, M. Bawa and P. Raghavan, *Heuristic-Based Address Clustering in Bitcoin* IEEE Access,November 2020.