Ledger Innovation Lab
03 rue Gretry, Paris
01/03/2022

Scope Statement:
Decentralized Test of Liveness

Calic Petar, Master ANDROIDE, Sorbonne Université
Supervised by: Nicolas Bacca, Co-Founder, VP Innovation Lab

SUMMARY

1/ Introduction

This project is to be realized by Petar Calic (me) under the supervision of Nicolas Bacca, VP of **Innovation lab at Ledger**, in the context of an internship of 6 months. I will be working mainly alone on the project with some assistance and guidance from Mr Bacca, the innovation team, the Data team, and many UPMC researchers that I will visit for consultations.
Know Your Customer (KYC) procedures are a critical function to assess customer risk and a legal requirement to comply with Anti-Money Laundering (AML) laws. Effective KYC involves knowing a customer's identity, their financial activities and the risk they pose.
Tests of liveness are a subgroup of tests designated to distinguish real and fake persons (bots). It consists of a set of technical features to counter biometric spoofing attacks where a replica imitates a person's biometrics or characteristics in order to deceive or bypass the identification and authentication steps given by the system. With the emergence of the internet with it, the daily increase of our everyday life dependence on its functionalities, increases the targets and vulnerabilities of third party malicious intentions. Here is one personal example: I made some nft's i wanted to give away freely to people that attended one online event i organized. The second I shared the link, a bot took all 100 of the tokens. A simple liveness check could prevent that. With AI capabilities to imitate humans rising every day, the task to discover bots that don't wish to be discovered, becomes more and more challenging.
However, there are still very successful tests like Google's CAPTCHA which relies on secret image recognition, mouse movements and a top secret server classifier. Our **goal** is to make an open source, completely decentralized Test of Liveness smart contract deployed on STARKNet, that would be accessible to anyone and if possible community driven. There is also a more profound and general mission behind the project which is to make advances in the field of decentralized technologies. We should start relying more on technologies that are verifiable than trustable.
An **important note** is that a test of liveness that relies purely on data collection and analysis, will never be 100% bot proof as the constant race of technologies between the two sides never stops (Not to mention the human farms for solving CAPTCHAs). In order to be completely sure that we are dealing with a real person is to request a document verification for really critical systems. However we are confident that this test will repel more than 90% of most advanced bots attacking the front of the program (Trying to solve the Test). The cost of breach will hardly be profitable.

2/ Image analysis

- Experiment with different techniques of Image processing : Noising, Denoising, rotations, cropping.
- Find a threshold where humans still easily recognise the image, and ML algorithms hesitate.
- Minimum 1000 images of different classes.
- Minimum 9 classes

3/ App development

The V1 will be a frontend Interface that will be optimized to work well on all platforms. It will request from the backend a 9 pictures (3x3) and a question and display it.
The pictures will be on a DB made by the Infrastructure team.
The test must be operational to begin V2
The V2 will have the DB of pictures stored on IPFS, and the back will operate on starkent via a SmartContract that will interact with the Front end interface. Chainlink for the random questions or a server(John says it's easier to begin with and cheaper).
The V3 will integrate the ML decider.

4/ Data Collection

Users that interact with the system on the V1 app will be used for building the dataset that will be used for classification training. The features need to be wisely chosen as we aren't going to be able to collect a lot of data and the information available is very restricted. We have to keep in mind that features that aren’t impacting the classification are prone to make the learning process overfitting. The data will be collected approximately with this pace: 10 tests for 20 users daily → 1000 data vectors in 5 days. With 500 we can begin the cross validation tests.

For bot data we should synthesize data. We will need slightly more data because the variations of behavior are greater than for a human. Web tech study or GAN.

5/ Classifier training

Choosing a simple but effective classifier. Make synthetic responses using a GAN. More cases need to be trained: an attacking ML that sees the answers, that partially sees, that is completely blind. Or maybe choosing a classifier that would make it as hard as possible exploiting the fact that the trained classifier is open source. For example using LSTM or Cascades or Haar.

8/ Cryptographic protection of labels

The system has a vulnerability which can be exploited with medium effort if the attacker has knowledge in Machine Learning. Otherwise the test of liveness itself is very robust and we are confident that it is very hard to breach the test using individually made ML bots that recognize Objects.
The attacker needs two elements in order to breach the system: Image labels (answers to the test), and the ML classifier. If retrieved the user can train it to find a set of features which satisfies the classifier and together with visible image labels the attacker can easily “act” as a human. If we find a way to protect one or those two elements the test will be almost 100% bot proof.

As we are focusing in this project on the decentralization object, no compromise will be made and centralize some operations to make the system more secure. The V3 will be opretational and robust against front attacks. The ML Open source vulnerability will exist but a maximum will be done to make the ML inference as complicated as possible

9/ Global architecture

Global

10/ Final Tests and Updates

Test

11/Economic Study

Compare Cost of breach, cost of using human farms, and price of the protected asset.

12/ Bibliographie

[1]