EXTENDED ABSTRACT: USING ARTIFICIAL INTELLIGENCE TO DETECT FRAUD ON THE BLOCKCHAIN

Ghassemi Toosi, Farshad, Lero, University of Limerick, Limerick, Ireland, farshad.toosi@lero.ie

Buckley, Jim, University of Limerick, Limerick, Ireland, jim.buckley@ul.ie

Sai, Ashish Rajendra, Lero, University of Limerick, Limerick, Ireland, ashish.sai@lero.ie

Le Gear, Andrew, Horizon Globex Ltd., Limerick, Ireland, andrew.legear@horizon-globex.ie

Context

Cryptocurrencies witnessed extensive attention from both academia and industry after the inflation in 2017. This broad recognition has led to its adoption by people beyond the research domain. This increased adoption may be attributed to the overall increase in the valuation of the cryptocurrencies which at its peak in 2017 reached a combined valuation of 900 Billion USD (*Cryptocurrency Market Capitalizations* 2019). At the time of writing, the cryptocurrencies hold a combined valuation of 200 Billion USD (*Cryptocurrency Market Capitalizations* 2019). This market capitalization makes cryptocurrencies very lucrative for attackers with malicious intents. A well-orchestrated attack on these cryptocurrencies may allow an attacker to attain monetary gain. The research on the security of blockchain has identified a number of attack vectors (Chia et al., 2018). One such attack vector is the coordinated manipulation of the blockchain services (for example, a pump and dump scheme to artificially inflate the price of an Initial Coin Offering (Li, Shin, and Wang, 2018)).

Objective

We propose using machine learning to identify the coordinated behavior of malicious entities in the blockchain network. We intend on training a neural network to identify patterns such as coordinated manipulation of the transactions to artificially inflate the hype around an Initial Coin Offering (ICO). For example, during an ICO, the company behind the ICO might try to create the impression that the offering has the potential to increase in value dramatically. They may do this by using a Hierarchical Deterministic (HD) wallet to seed new private keys, creating the impression that each purchase on the ICO comes from a different wallet/investor. This suggests that interest in the offering is of more widespread appeal than would otherwise be perceived, and may prompt other investors to purchase, resulting in an increase in the ICO's value (Known as 'pump-and-dump' fraud). Some work has already been undertaken (Treleaven, Batrinca, et al., 2017), (Dinh and Thai, 2018), (Marwala and Xing, 2018), and the work of our group builds on this existing research by investigating and leveraging approaches from the field of machine learning, towards detection of pattern-based fraud on the blockchain.

Approach

Machine learning as an application of artificial intelligence (AI) provides a model that automatically learns from some trustable information to predict some unseen (future) data according to what was learned. The neural network is one of the several different techniques in Machine learning that comprised of a set of neurons in different layers in which the neurons of each layer interact with the neuron of the immediately next layer via some synapse (Hinton, Vinyals, and Dean, 2015). The minimum number of layers that a neural network can have is two where the first layer accommodates the input data and the second layer accommodates the output data. The number of neurons at the first and the last layers of each neural network indicates the dimensionality of the input and output data respectively, e.g., a neural network with n neurons at the first layer indicates that the input data has n features (dimensions).
Transactions in blockchain and cryptocurrency often reveal complicated and at the same time meaningful patterns (e.g., 'pump-and-dump'). For instance, initiating a coin offering (known as ICO) and performing some intended and cyclic transactions on the initial coin offering smart contract in order to draw attention to increase the price of the coin is considered as a fraud ((Li, Shin, and Wang, 2018)). To bring this scenario into machine learning (more specifically neural network) context for fraud detection one needs to define what is input and output. Our idea to create such a neural network, is to create a numerical model of the ICO transactions and use it as the input data to

feed the input layer of the network. The decision whether this set of ICO transactions is considered as a fraud or not is the output layer of the neural network.
There are different types of neural networks such as *Perceptron*, *Linear Regression*, *Logistic regression* and *nonlinear Regression*. Depends on the format of input and output data one of the aforementioned neural networks is chosen. In this work we are going to use a *Linear Regression* to train some known data to predict some unknown data. We are using *Linear Regression* as the format of the output is not defined as a binary value (e.g., fraud and ordinary). We define a strength for the fraud; for instance, a set of ICO's transactions might show around 70% fraud.

Illustration

To illustrate this approach, we gather some sets of transactions with their status (e.g., the strength of the fraud) and divide them into two different sets, *train* and *test* data. We train the model using the *train* set and test our model against the *test* set. In order to eliminate the negative effect by arbitrary distribution of data into *train* and *test* sets, we apply cross-validation technique (Pedregosa et al., 2011) and calculate the *Ein* (error on training set) and *Eout* (error on test set) to validate the model Abu-Mostafa, Magdon-Ismail, and Lin, 2012.
The first step of constructing the model is to convert the set of transactions into some readable format for the neural network. A suitable type of data for Neural network is numerical data. Data is usually presented either as numerical data, categorical data or relational data. In the case of cryptocurrency, we are dealing with relational data. Therefore, the first task is to convert the extracted relational data into some numerical data suitable to be fed into the neural network. The second step is to estimate the strength of the fraud for the given set of transactions as output data (values of the neurons at the last layer). The next step would include running the network to optimize the weights of the links between layers (constructing the model). To do this, we make use of a well-known Machine learning Python library called *scikit-learn* (Pedregosa et al., 2011). Once the network is trained, the test data is fed into the network and the generated output is compared against the ideal output and error is calculated (*Eout*). The consider a neural network as an ideal neural network when both *Ein* and *Eout* are small and close to each other.

Conclusion

Our initial investigation shows how machine learning can be used to identify pattern based frauds on the blockchain. For our initial analysis, we investigate a Pump and Dump scheme and propose using machine learning techniques such as *Linear Regression* for the detection. We also propose using a strength function to categorize an ICO as fraud rather than using a binary classification.
Another possible use of machine learning for fraud detection is in the identification of wallets with a high degree of cohesiveness. By detecting wallets with high cohesiveness, we can identify the source of potential fraud.
Conclusively, we have illustrated the advantage of machine learning techniques when practiced on the blockchain data. Given the abundance of existent machine learning techniques, we plan to do an extensive review of these techniques and the data accommodated in the blockchain, to classify possible investigation techniques and analysis targets. We then plan to coordinate these targets to the suitable techniques and appraise the effectiveness of these techniques in this novel context.

References

Abu-Mostafa, Y. S., M. Magdon-Ismail, and H.-T. Lin (2012). *Learning from data*. Vol. 4. AMLBook New York, NY, USA:

Chia, V., P. Hartel, Q. Hum, S. Ma, G. Piliouras, D. Reijsbergen, M. van Staalduinen, and P. Szalachowski (2018). "Rethinking blockchain security: Position paper." *arXiv preprint arXiv:1806.04358*.

Cryptocurrency Market Capitalizations (2019). URL: `https://coinmarketcap.com/`.

Dinh, T. N. and M. T. Thai (2018). "Ai and blockchain: A disruptive integration." *Computer* 51 (9), 48–53.

Hinton, G., O. Vinyals, and J. Dean (2015). "Distilling the knowledge in a neural network." *arXiv preprint arXiv:1503.02531*.

Li, T., D. Shin, and B. Wang (2018). "Cryptocurrency pump-and-dump schemes." *Available at SSRN*.

Marwala, T. and B. Xing (2018). "Blockchain and artificial intelligence." *arXiv preprint arXiv:1802.04451*.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. (2011). "Scikit-learn: Machine learning in Python." *Journal of machine learning research* 12 (Oct), 2825–2830.

Treleaven, P., B. Batrinca, et al. (2017). "Algorithmic regulation: Automating financial compliance monitoring and regulation using ai and blockchain." *Journal of Financial Transformation* 45, 14–21.