CONTINUOUS SECURITY IN IOT USING BLOCKCHAIN

Rahul Agrawal, Pratik Verma, Rahul Sonanis, Umang Goel, Dr. Aloknath De, Sai Anirudh Kondaveeti, Suman Shekhar

Samsung R&D Institute India, Bengaluru

Summary

- Lack of trust between various entities in Internet of Things (IoT) and single point of failure are two major roadblocks. Using immutable and decentralized nature of Blockchain [1], IoT system can be made more robust establishing trust and avoiding single point of failure.
- In this work, we demonstrate blockchain based IoT solution in Smart building scenario, establishing continuous security by validating user trail at every interaction.
- Unique crypto-tokens are required for transaction in blockchain which are pre-generated using prediction models based on user trail.

IoT-Zone Identification

IoT-Zone Identification

- Initial topology of IoT-zone is established based on physical connections (like swipe gates) between multiple zones and set of rules as shown in Fig 1.
- User's transition in zones is modeled as a random variable whose state-transitions can be represented via a directed graph with edges having respective state transition probabilities.

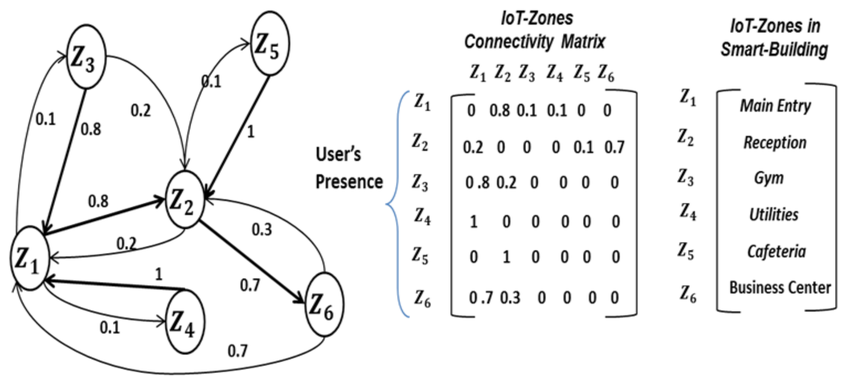

Fig 1: Rule Based IoT-Zone Connectivity

- Variable Markov Model [2] and LSTM based RNN prediction models [3] are used to predict user's next possible zone visit.
- A threshold Markov model of order n is chosen. Let the next zone (i.e. state) be denoted by Z_{n+1}. Let $\mathbf{z} \equiv Z_1 ... Z_n$ be the previous sequence of up to n zones. Let $N(\mathbf{z}Z_{n+1})$ denote the number of occurrences of the subsequence of zones encountered in the training sequence. Let

$$\Omega = \{\sigma : N(\mathbf{z}\sigma) > 0\}$$

then, the conditional probability estimator is given by

$$\widehat{P_n}(Z_{n+1}|\mathbf{z}) = \frac{N(\mathbf{z}Z_{n+1})}{\sum_{Z \in \Omega} N(\mathbf{z}Z)}$$

- For implementing the n bounded Markov model, data-structure "*trie*" $\boldsymbol{T}$ is used. Each node of $\boldsymbol{T}$ represents a zone and has a counter for number of visits made so far. Initially, during the training phase, the "*trie*" is initialized with valid nodes based on history of user's trail.

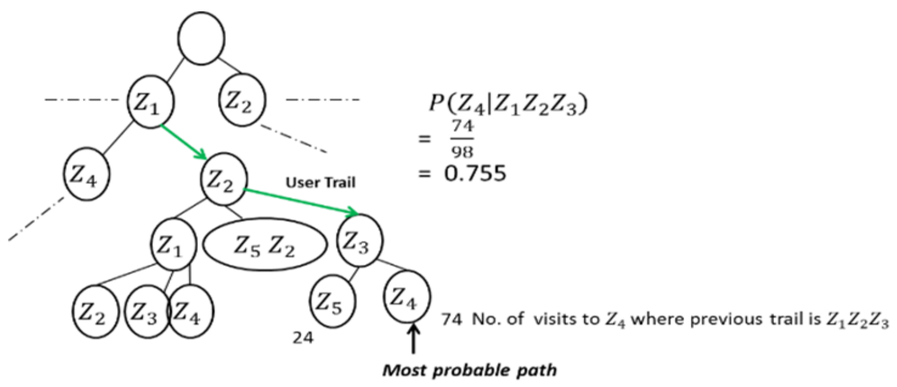

Fig 2: A "*trie*" corresponding to a sample third order model

IoT-Token

IoT-Token Generation

- Every user in IoT system has to register with Certificate Authority by providing his attributes like permission level, organization etc. for generation of user's private-public key pair.
- User's public key is used to gather transactions from blockchain establishing his/her trail. Using prediction model, IoT-tokens are generated for high probability zones whereas for low probability zones, a second factor of authentication is needed.

```
{
        "timestamp": "1572042",
        "userName": "Charlie",
        "currentLocation": "Floor 4",
        "userTrail": ["Parking", "Entrance", "Cafe"]
        "designation": "employee",
        "ECert": "MIIBrjCC....CQ"
}
```

IoT-Token Validation

- User triggers an interaction which is analyzed by nearby IoT-hub and corresponding IoT-token is queried from digital wallet of nearby user devices.
- Token queried carries a digital signature which is verified by the network using user public key to ensure token is not used by other user in case of token theft.
- If token is authentic, token is included as a part of transaction which is added in blockchain using Practical Byzantine Fault Tolerance Consensus (PBFT).

Experiments

- Blockchain network framework used is Hyperledger Fabric v0.6 with 4 validating nodes and 1 certificate authority node.
- Dataset of multiple users transitioning between zones in a smart building is used each comprising of 1000 data sequence. Gradient descent algorithm is used for optimization in LSTM model.

Results

- Table 1 presents the prediction accuracy for different users for both LSTM model and 3rd order Markov Model.
- Figure 3 represents the zone connectivity graph generated using accumulated data of every user in the system.

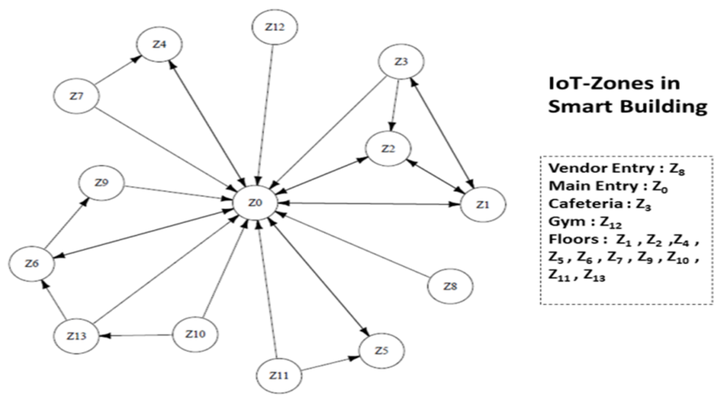

Fig 3: Rule Based IoT-Zone Connectivity

	Markov			LSTM	
User Id	3^{rd} order	2^{nd} order	1^{st} order	No of labels	Accuracy
1	98.75	99.12	99.30	4	77.77
2	89.65	89.65	93.10	4	71.80
3	43.75	59.37	62.50	5	64.40
4	85.18	88.89	96.30	6	61.50
5	97.88	83.10	66.90	7	60.00
6	53.33	63.33	73.33	5	52.68
7	54.76	64.29	73.81	7	52.08
8	61.19	64.18	65.67	7	46.85

Table 1: Accuracy of Markov and LSTM model for 8 users

Discussions

It is observed that, in case of Variable Markov Model, decreasing the order improves accuracy with a majority of the users, suggesting that for the user zone movement prediction, the immediate history has more correlation with the next step. In case of LSTM model, on an average, it gives less prediction accuracy than Markov model, owing to lesser training data and the inherent first order Markov behavior present in the training set.
Our approach enhances the security of IoT system without any user intervention as crypto-tokens are pre-generated using various prediction models. Further crypto-token generation can be improvised by using an ensemble learning approach which uses the best models specific to the input data or a weighted combination thereof.

Contact

Pratik Verma
Samsung R& D Institute, Bangalore
pratik.verma@samsung.com

References

[1] K. Christidis and M. Devetsikiotis, "Blockchains and Smart Contracts for the Internet of Things," IEEE Access, vol. 4, pp. 2292-2303, 2016.

[2] Ron Begleiter, Ran El-Yaniv, and Golan Yona, "On Prediction Using Variable Order Markov Models," Journal of Artificial Intelligence Research, vol. 22, no. 1, pp. 385-421, Dec. 2004.

[3] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2222-2232, Oct 2017.