CHITTAGONG UNIVERSITY OF ENGINEERING & TECHNOLOGY (CUET)
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
CHITTAGONG – 4349

(Project Proposal)

Application for the Approval of B.Sc. Engg. Thesis / Project
(Computer Science & Engineering)

Date: 23.09. 2019

1. **Name of Student** : Md. Hafizul Haque
   Roll No. : 1504097 **Session:** 2018-2019

2. **Present Address** : A-410, Bangabandhu Hall, Chittagong University
   of Engineering & Technology, Raozan, Chittagong

3. **Name of Supervisor** : Dr. Md. Mokammel Haque
   Designation : Professor
   Department of Computer Science & Engineering (CSE)
   Chittagong University of Engineering & Technology

4. **Name of Department** : Computer Science & Engineering(CSE)

   Program : B.Sc. Engineering

5. **Date of First Enrollment**
   in the program : February 25, 2016

6. **Tentative Title** : Blockchain-Based Malware Detection and Evidence
   Extraction in Android Devices.

7. Introduction:

Blockchain is a technology that can best be thought of as a ledger that maintains a record of economic transactions or anything of value. The ‘block’ element of blockchain refers to the fact that groups of transactions are first batched to form a block, and are then combined to form a chain, hence the term blockchain. One characteristic about blockchain that makes it valuable as a piece of technology is that it is tamper-resistant against bad actors. Thus, once data is batched into blocks and then added to the chain, it becomes extremely difficult to change the data included within those blocks. This tamper-resistance derives from the fact that, included in every block is a cryptographic hash of the previous block. As more blocks are added to the chain, going back to alter data within a previous block would require a bad actor to recompute the hash of that block and all blocks after it, a task that would come at a considerably high financial cost [1].

A blockchain is usually managed or over-seen by a second-layer network of peer-to-peer computing nodes. These nodes can be thought of as computers that run a specific client that allows them to connect to the blockchain. Nodes are effectively computers that represent an individual or entity on a blockchain based network. Each node is permitted a full copy of the blockchain, upon joining the network. These peer-to-peer computing nodes are vital because they validate and verify all transactions that are made on the network. Nodes that validate transactions are referred to as ‘mining nodes’ and they secure the network by ensuring that only correctly formed transactions and blocks are included on the blockchain. This process is known as reaching consensus, and consensus algorithms such as: proof-of-work, proof-of-stake, delegated proof-of-stake, proof-of-importance and more are utilized to help facilitate this process.[2]

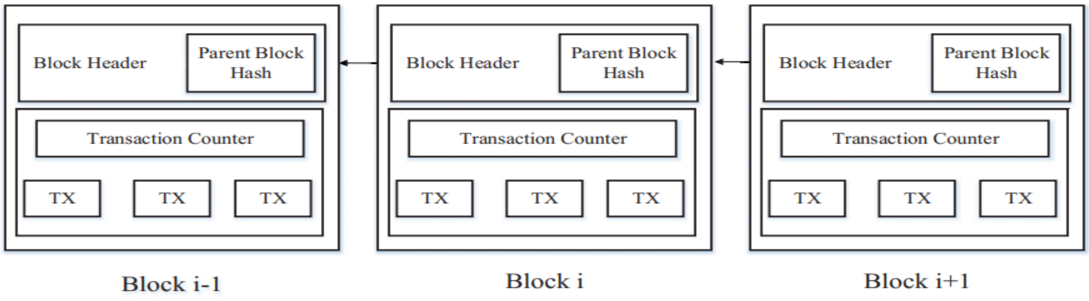


Figure 1:An example of blockchain which consists of a continuous sequence of blocks

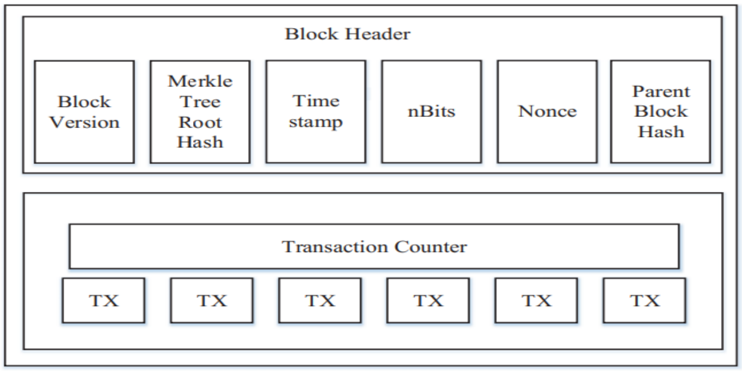


Figure 2: Block Structure

Currently, there are at least four types of blockchain networks — public blockchains, private blockchains, consortium blockchains and hybrid blockchains[2]. We will use mixed chains consisting a consortium chain and a public chain.

Consortium blockchains differ to their public counterpart in that they are permissioned, thus, not just anyone with an internet connection could gain access to a consortium blockchain. These types of blockchains could also be described as being semi-decentralized. Control over a consortium blockchain is not granted to a single entity, but rather a group of approved individuals. With a consortium blockchain, the consensus process is likely to differ to that of a public blockchain. Instead of anyone being able to partake in the procedure, consensus participants of a consortium blockchain are likely to be a group of pre-approved nodes on the network [1]. Thus, consortium blockchains possess the security features that are inherent in public blockchains, whilst also allowing for a greater degree of control over the network.

8. Background and Present State of the Problem:

It is always a hot issue for the related security problem in mobile devices and wireless network, which has been studied by many researchers. Nowadays, one of the typical methods in malware detection is the feature extraction like signature and permission information. Some of them are described bellow:

For instance, the Ensemble Learning (EL) extended the feature set of detection and proposed the classification way of 179 features, including API calls, instructions and authorities for detecting zero-day Android malicious code [3].

Another one is ANNCMDroid. The mobile malware detection method ANNCMDroid was based on co-occurrence matrices and artificial neural networks, which took into account the relations among sequences of system call [4].

The static stain analysis tool—DidFail was developed by combining the inter-component communication detection engines FlowDroid and Epicc [5].

The concept of security detection in andoird devices using blockchain is somewhat new. Generally, the malware detecting technologies for Android devices can be divided into static-based analysis and dynamic-based analysis. The static-based analysis method can implement efficient and automatic analysis in some ways. However, it cannot detect code obfuscation and encryption, and it is insufficient to decrypt malicious code in dynamic execution. Also, it uses a coarse-grained detecting approach of information flows between applications, which is easy to produce false positive [6]. The dynamic-based analysis method of Android-based software collects applications' behavior information during its operation. They can solve the problems such as code obfuscation and encryption. Nevertheless, malware usually have a well-designed triggering mechanism when facing dynamic tests, while some malicious programs can detect their own operating environments and automatic crash behaviors when running under the simulator [6]. In addition, using a single feature to determine software's malice is far below satisfaction. At the same time, the extraction of some features in the existing methods requires high time cost. For example, the extraction of API context with Appcontext requires a lot of time and memory, and its experiment just dealt with samples whose software packages were less than 5 MB. Recently, the blockchain technology has gained much more focus. Its key technology derives from the consensus mechanism, which is an example of a distributed computing system with high fault tolerance. The common technologies are: Proof of Work (PoW), Proof of Stake (PoS), Practical Byzantine Fault Tolerance (PBFT), Delegated Proof of Stake (DPOS) and so on. Generally, the research of the blockchain technology can be divided into three categories: the public blockchain, the private blockchain and the consortium blockchain. From the public blockchain, anyone can read and send the transaction, which can be effectively recognized, and anyone can participate in the consensus process. Thus, in the public blockchain, all records are visible to the public and everyone can participate in the consensus process. In the private blockchain, only those users from specific organizations can be allowed to participate in the consensus process. In the consortium blockchain, the consensus process is controlled by preselected nodes, which maintain a copy of the distributed data store. That is, only a set of preselected users can participate in the consensus process [7].

The consortium blockchain is a community of *N* member organizations, each of which runs a node. And in order for each block to take effect, it requires the confirmation of 2/3 of the organizations [8]. Thanks to the flexibility, the consortium blockchain technology has been applied to both financial and non-financial systems .

9. Objective with Specific Aims and Possible Outcomes:

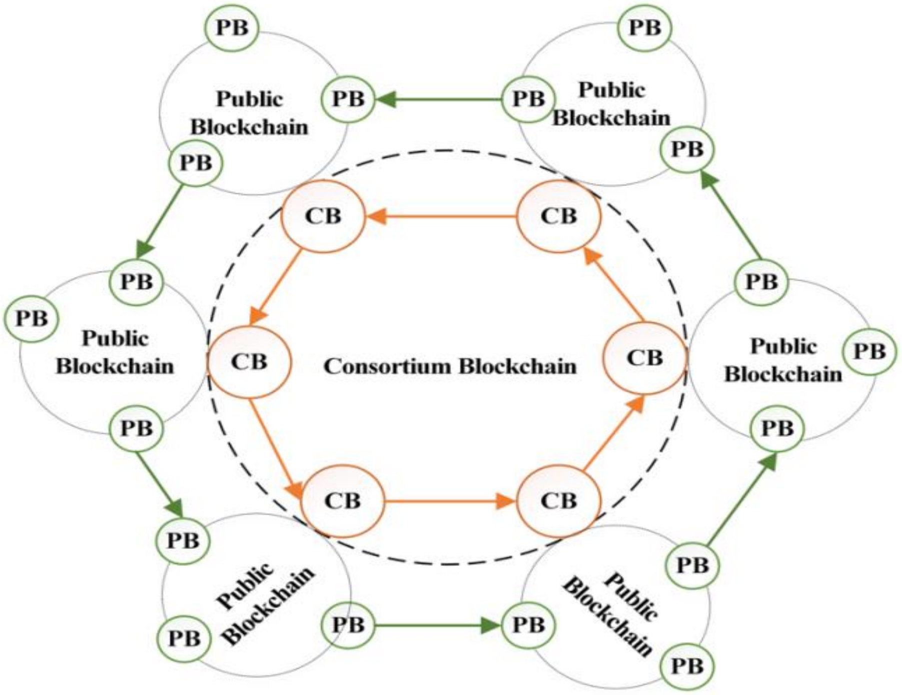
The main objective of my work is illustrated in the following:

- We propose a framework of Blockchain for Malware Detection and Evidence Extraction (BB-MDEE) in Android devices. The framework is composed of two parts of mixed chains: detecting consortium chain by test members and public chain by users.
- We analyze different malware families on the basis of Android-based systems and build a corresponding Multi-Feature Model (*MFM*).
- To reduce false-positive rate and improve the detecting ability of malware variants, we propose multiple marking functions. From this model, we can extract the features to construct the feature database, and develop a multi-feature detection algorithm.

10. Outline of Methodology/Experimental design:

10.1 Structure of BB-MDEE framework:

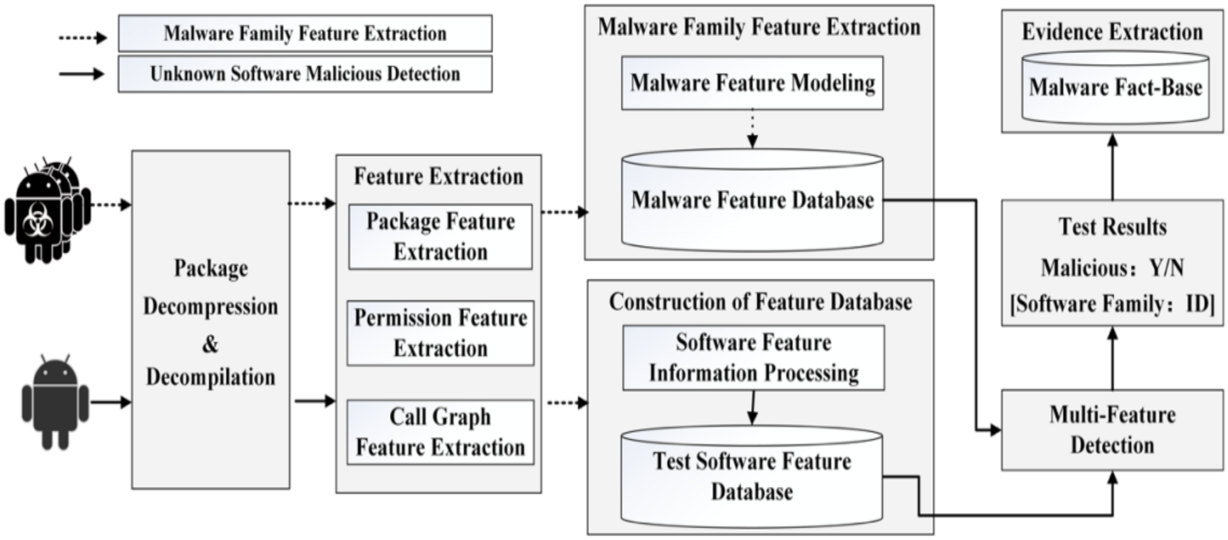
Our BB-MDEE framework consists of consortium blockchain and public blockchain, as shown in figure bellow, where CB represents the Consortium Blockchain and PB represents the Public Blockchain.



The CB is the core chain, composed of the members in distributed malware detection organizations. These members build a fact-base of the distributed malicious codes. The PB is the application chain, open for any user who needs to provide detection and evidence services for joining as a member node.

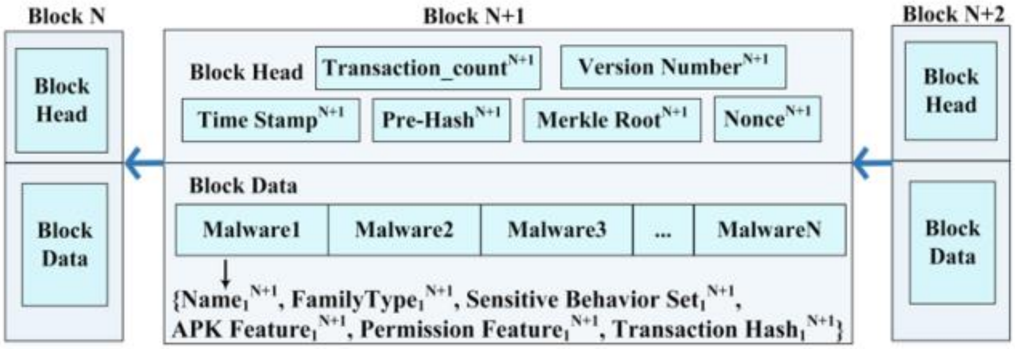
10.2 Multi-Feature Detection Process:

Our malware detection process is capable of both extract malware family features and detect unknown malicious applications. For detecting malware it follows *tiant analysis* method. For both malicious and unknown type of android application it first disassemble and decompile the packages. Then it extracts the features of application such as package, permission, call graph etc. Then it compares those features with malware feature database. It also processes the extracted features and then test them against the test software feature databases to find anomaly. After then it applies multi-feature detection algorithm and takes decision whether the application is malicious or not. if it is proved to be malicious then it forms malware fact base. And later these fact bases are stored in a new block in the core consortium blockchain.


Malware Family Feature Extraction
Unknown Software Malicious Detection
Package Decompression & Decompilation
Feature Extraction
Package Feature Extraction
Permission Feature Extraction
Call Graph Feature Extraction
Malware Family Feature Extraction
Malware Feature Modeling
Malware Feature Database
Construction of Feature Database
Software Feature Information Processing
Test Software Feature Database
Evidence Extraction
Malware Fact-Base
Test Results
Malicious：Y/N
[Software Family：ID]
Multi-Feature Detection

10.3 Constructing blocks in the chain:

Malware fact bases contain information like name, family type, sensitive behavior set, apk feature, permission feature, transaction hash. A demo block in the consortium chain that contains malware fact bases in it’s body is shown in the following figure:



11. Resources Required Completing the Work

The resources that are required by our proposed system are specified below:

- Personal Computer.
- Operating System windows 10.
- FlowDroid Tool for data flow analysis in Android

12. Cost Estimate

The costs that will occur to implement out proposed system are given below:

a) Cost of Materials:

Personal Computer (PC)	Tk. 60,000.00
Paper Cost	Tk. 300.00
Total	Tk. 60,300.00

b) Typing, Drafting, Binding:

Internet Browsing and Typing	Tk. 1,000.00
Drafting	Tk. 200.00
Binding	Tk. 200.00
	Tk. 1,400.00
Total	Tk. 61,700.00

13. References:

[1] B. Asolo, “Consortium Blockchain Explained.” [Online]. Available: https://www.mycryptopedia.com/consortium-blockchain-explained/. [Accessed: 21-Nov-2019].

[2] D. S. Michael Nofer, Peter Gomber, Oliver Hinz, “Blockchain,” *Springer*, no. March, pp. 1–6, 2017.

[3] S. Y. Yerima *et al.*, “High Accuracy Android Malware Detection Using Ensemble Learning,” *IET Inf. Secur.*, 2015.

[4] X. Xiao, Z. Wang, Q. Li, Q. Li, and Y. Jiang, “ANNs on Co-occurrence Matrices for Mobile Malware Detection,” *KSII Trans. INTERNET Inf. Syst.*, vol. 9, no. 7, pp. 2736–2754, 2015.

[5] and W. S. J. Burket, L. Flynn, W. Klieber, J. Lim, W. Shen, “Making DidFail succeed: Enhancing the CERT static taint analyzer for Android app sets,” *Softw. Eng. Inst., Washington, DC, USA, Tech. Rep. C. 2015-TR-001*, 2015.

[6] B. Amro, “MALWARE DETECTION TECHNIQUES FOR MOBILE DEVICES,” *IJMNCT*, vol. 7, no. 4, pp. 1–10, 2017.

[7] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, “An Overview of Blockchain Technology : Architecture , Consensus , and Future Trends,” *IEEE*, 2017.

[8] Y. Z. A. Z. W. JINGJING GU, BINGLIN SUN, XIAOJIANG DU, JUN WANG, “Consortium Blockchain-Based Malware Detection in Mobile Devices,” *IEEE Access*, vol. 6, pp. 12118–12128, 2018.

14. CSE Undergraduate Student (CUGS) Committee reference:

Meeting No: **Resolution No:** **Date:**

15. Number of Under-Graduate Student(s) working with the Supervisor at Present: 09

Signature of the student

Signature of the Supervisor

Signature of the Head of the Department