Towards Reliable Integrity in Blacklisting: Facing Malicious IPs in GHOST Smart Contracts

Georgios Spathoulas[1*], Anastasija Collen[2], Pankaj Pandey[1], Niels A. Nijdam[2], Sokratis Katsikas[1],
Charalampos S. Kouzinopoulos[3], Maher Ben Moussa[2], Konstantinos M. Giannoutakis[3],
Konstantinos Votis[3], Dimitrios Tzovaras[3]
[1]*Center for Cyber and Information Security, Norwegian University of Science and Technology*, Gjøvik, Norway
[2]*Centre Universitaire d'Informatique, University of Geneva*, Geneva, Switzerland
[3]*Information Technologies Institute, Centre for Research and Technology Hellas*, Thessaloniki, Greece
[*]gspathoulas@dib.uth.gr, [2]{anastasija.collen,niels.nijdam,maher.benmoussa}@unige.ch,
[1]{pankaj.pandey,sokratis.katsikas}@ntnu.no, [3]{kouzinopoulos,kgiannou,kvotis,dimitrios.tzovaras}@iti.gr

Abstract—**When it comes to novel cyber security solutions for extremely heterogeneous environments like Internet of Things (IoT) and smart homes, the key focus is typically given to the understanding of network activities and elimination of suspicious traffic. The GHOST project adds an extra dimension to this approach by integrating blockchain technology at its core decision mechanism. On a daily basis, each GHOST installation is encountering malicious behaviour and suspicious IoT communications, where easy information sharing with other installations, as well as decentralised decision making, are mandatory features for the efficient protection of the end-user. GHOST's Smart Contracts (SC) are designed to tackle in an easy, yet productive way, the reporting on suspicious IP addresses which the IoT devices in a smart home are trying to communicate with. Two variations of blacklisting smart contracts are presented in this paper, covering a diverse spectrum of possible attack vectors while closely following the Privacy by Design (PbD) principles. A reputation scoring scheme for malicious IPs reporting is integrated in the SC, uncovering the implementation details on the penalisation of existing entries in case of malicious behaviour of reporting devices.**

Index Terms—**blockchain, smart contracts, blacklisting, privacy, cyber security, information sharing, smart homes**

I. Introduction

Traditional cyber security solutions mostly offer server-centric design, causing difficulties in decision making distribution and information sharing. Additionally, such setups are prone to a system's single point of failure. The use of the blockchain technology exploits the true benefits of the decentralised approach, by replicating valuable data on each network's node, thus enforcing dynamism and reliability of the whole system.

Additionally, in the context of smart home security, various problems arise from the unmonitored and unknown data exchange of IoT devices that are installed in smart home solutions, with external parties.

The European research project GHOST (https://www.ghost-iot.eu/) challenges the traditional cyber security solutions for the IoT sector by proposing a novel reference architecture that is embedded in an adequately adapted smart home network gateway, and is designed to be vendor-independent. It proposes to lead a paradigm shift in consumer cyber security by coupling usable security with transparency and behavioural engineering. The GHOST project is exploiting novel technologies, such as blockchain, to provide resilience and integrity of decision making on the communication exchange.

In particular, the blockchain technology integrated in GHOST aids to counteract the most modern attacks by decentralising its core risk assessment decision making engine, *Risk Engine (RE)*. The SC assimilated in GHOST aim to facilitate data sharing on the malicious IPs reported by individual installations and retrieved from the online collective intelligence, *Cyber Security Knowledge Base (CSKB)*. The detailed architectural setup and overview of the complete blockchain integration into the cyber security solution are discussed in our previous works [1], [2].

This paper is focusing on one specific SC use case, that aims to provide decentralised security for the protection against data exchange with malicious nodes external to the smart home. Each GHOST smart home installation will collaboratively create and maintain a blacklist of malicious IP addresses, by sharing RE produced data from the evaluation of the risks imposed by specific connections between the external IPs and the gateway, an external IP and the IoT devices in the smart home, and the actual behaviour of the IoT devices and their network communication profiles. Based on the calculated risk metrics, the GHOST system can characterise external IPs based on the probability that they are malicious. This information will be stored on the private blockchain with the help of smart contracts. The implementation of the blacklist implies operation with data of different levels of sensitivity and trustworthiness and, therefore, can be split into two scenarios: public and private.

The rest of the paper is structured as follows. Section II outlines the research method followed for producing the results presented in this paper. Related work is described in Section III. Section IV presents the Blacklisting IP use case that was developed within the GHOST project, and open issues are discussed in Section V. Our conclusions are summarised in Section VI.

Design Science Activity	**Corresponding Section in this Paper**
Explicate Problem	Introduction
Design Requirements	Introduction and Related Work
Design and Development of Artefact	Use case description
Demonstration of Artefact	Implementation
Evaluation of Artefact	Implementation

TABLE I
DSRM AND PAPER'S CORRELATION

II. RESEARCH METHOD

The Design Science Research Method (DSRM) [3] is followed in this paper. Design science research relies on the creation of "knowledge and understanding of a design problem, and its solution is acquired in the building and application of an artefact" [4]. DSRM also involves "analysis of the use and performance of designed artefacts to understand, explain and very frequently to improve the behaviour of aspects of information systems" [5]. Johannesson and Perjons [3] presented a DSRM Framework consisting of five main activities as shown in Figure 1. These activities and their coverage by corresponding paper sections are presented in Table I.

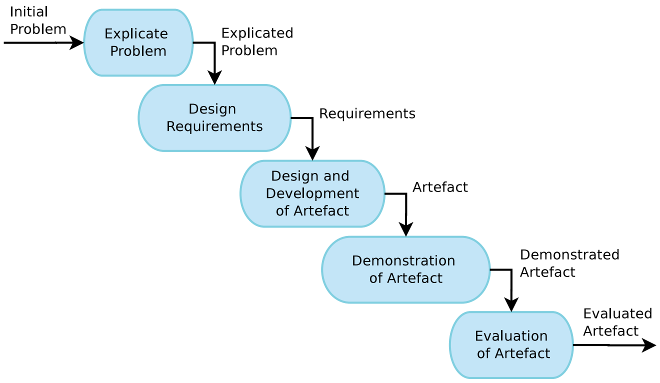

Fig. 1. Design Science Research Method

Offermann et al. [6] studied the artefacts designed and developed in the area of information technology and information systems, and classified the artefacts into eight categories: System Design, Method, Language/Notation, Algorithm, Guideline, Requirements, Pattern, and Metric. According to that classification of artefacts, this paper presents a "Method - Defines the activities to create or interact with a system" and "Algorithm - An executable description of behaviour of a system" types of artefacts.

Vaishnavi and Kuechler [5] presented a set of methods for evaluation and validation of design science artefacts as shown in Table II. This paper adopts a mix of "Demonstration" and "Logical Reasoning" forms of evaluation according to this classification of evaluation (validation) methodologies.

Method	**Description**
Demonstration	The demonstration method is the weakest form of validation. This method is appropriate for the solutions that are novel and solve a problem for which no other solution exists.
Logical Reasoning	The strength of the logical reasoning form of evaluation depends on the strength and preciseness of the arguments and assumptions. This form of evaluation is an alternate method for experiment and simulation methods.
Experiment and Simulation	This method is useful when the problem is complex and not receptive to a mathematical proof.
Using Metrics	The use of metrics is useful in experiments, simulation, and mathematical proof methods. It is valuable in quantification of the claims about the artefact.
Benchmarking	The benchmarking evaluation method is a weaker form of using metrics method. This method is useful in comparing the results obtained from the experiments and simulations method. This method is useful when there is a lack of suitable metrics to validate the artefact's claims.
Mathematical Proof	The mathematical proof form of evaluation is the strongest form of validation.

TABLE II
DESIGN SCIENCE EVALUATION METHODS

III. RELATED WORK

The blockchain technology has been recently applied to multiple domains to test its efficiency and to evaluate whether it can be an adequate solution to various problems. One such domain is the IoT. The main reason behind this choice is the structure of the IoT, that is currently evolving to the Internet of Everything, where numerous devices are going to be interconnected and will interact with each other [7], [8]. These devices are going to be mostly autonomous and they shall communicate on an equal basis, without any central node to provide the required trust. Using the blockchain technology, which mainly provides trust between nodes, seems to be an effective approach in order to facilitate the future underlying infrastructure for IoT. On the other hand, serious limitations hinder such implementations, as the main approach for consensus, proof of work, demands significant processing resources, which are probably non existent in the IoT ecosystem. Some of the most interesting approaches to the combination of blockchain technology and the IoT paradigm are discussed in this subsection. Huckle and others present scenarios in which IoT and blockchain can be used in order to enable sharing economies of different assets [9]. One issue with applying blockchain in sharing economy applications is the interaction of the assets with the blockchain infrastructure. By using smart IoT devices it is possible to automatically restrict or grant access to assets like vehicles or buildings according to rules implied by smart contracts without the need for any human intervention. Another important problem regarding IoT is security and one of the factors making this problem worse is the fact that IoT devices often function with outdated firmware. Lee and Lee propose to use the blockchain to certify IoT devices running on the latest and most secure firmware [10]. This is an interesting approach, that could also employ the creation of an open market between manufacturers, end-users, validators and penetration testers. However, there are limitations with respect to the application of the proper rules in the devices to provide the initiative for all stakeholders to push for more secure firmware installed on deployed devices. Supply chain monitoring has been one of the first domains for the application of blockchain technology. Specifically, blockchain allows to monitor the whole process of producing an end product from the moment when the initial ingredients are produced, through the whole manufacturing and exchange process and until the moment the end user buys the product. This approach can have a significant effect on the quality of products sold. The IoT devices are needed to monitor the process in the most secure and non intrusive way [11].

The use of the blockchain technology also recently gained interest in the cyber security domain. In particular, the use of smart contracts was incorporated in the design and implementation of a DDoS defence mechanism in [12]. The use of the existing public infrastructure of Ethereum to advertise blacklisted IPs suspected to be involved in ongoing DDoS attacks is fully exploited in this work. However, despite the advantageous reuse of existing infrastructure and the addition of a novel security mechanism for easy data sharing, a necessity of a central element for proper system functioning still remains at its core. A smart contract *Registry* is required for issuing certificates of the ownership of IP addresses, when submitting new data to the smart contracts. The use of timeout notions while creating white/black/grey lists of IP addresses in the light of the ChainGuard firewall functionality and nodes classification is discussed in [13]. Any node is eventually placed in the greylist, and on the timeout expiration or inactivity from the application itself, it is moved to the blacklist. Alternatively, the nodes are transitioned to the whitelist upon granted connection permission. The greylist capacity is fixed, and once it is full, the ChainGuard generates short-lived flow entry with instruction to drop similar packets. This eventually serves as a DDoS attack monitoring mechanism. Another interesting approach of using blacklists is discussed in [14], in the field of Vehicular Wireless Networks. When a node is accused of one or more malicious acts, the reporting party prepares a submission of the blacklist data on the blockchain network, including information on the vehicle unique identification, a timestamp of when the ban was issued, a timestamp of when it will expire, and the identities of the source and destination nodes, together with their public keys. These entries are stored on each node locally with distributed access ensured by a blockchain network.

There are some recent research attempts to use the blockchain technology in order to implement reputation or scoring systems for various use cases. Dennis and Owen propose a blockchain based reputation system [15]. After the interaction between two nodes, each node can commit a transaction consisting of the reputation score, a timestamp, and a hash of the interaction. The authors state that the proposed methodology may solve many problems of current reputation systems, such as unfair ratings attack, collusion attack, sybil attack and re-entry attack. Zhao and others discuss the use of blockchain in order to build a decentralised system capable of emulating the functioning of traditional publish-subscribe systems [16]. Publishers publish a topic on the blockchain and subscribers specify an interest message by making a deposit to subscribe to the topic. Then, if there is a match, the publisher transmits encrypted content to the blockchain, subscribers decrypt it, and mark the publisher as its reputation. Finally, the publisher receives the payment from the subscriber. An implementation of the protocol on Ethereum is also demonstrated in this work. Schaub and others describe a decentralised reputation scheme for e-commerce [17]. They use a blockchain structure to store ratings of service providers submitted by their customers. Each new rating is accompanied by a reference to the block of the previous rating for the specific provider, to enable fast summation of all ratings of a provider. Additionally, the authors propose the use of blind signatures to protect the anonymity of the customers, while ensuring that only the customers that should do so submit ratings.

IV. Smart Contract for Blacklisting IPs

The GHOST network will consist of a full-scale blockchain deployment: Smart Home installations, Smart Sensors and Ethereum nodes. Apart from the main, live network though, a second experimentation private Ethereum [18] blockchain network will be maintained during the development phase of the project. The experimentation blockchain network will act as a testbed for the development and testing of SC prior to their release on the live network.

Ethereum is probably the most open and flexible among established blockchain implementations. It enables the developer to construct smart contracts, which are more or less equivalent to normal traditional programs in terms of functionality, while they are executed in a completely decentralised way. An Ethereum network is constructed from multiple Ethereum nodes, which are running instances of one of the Ethereum clients. For the GHOST blockchain network Geth, the most popular client, will be used.

For the purposes of the GHOST project, a node has to be installed in each smart home. This node will be installed at the smart home gateway middleware. Because of limitations in hardware resources, these nodes are going to be light nodes and will not conduct mining. Additionally, in order to make the Ethereum blockchain functional, multiple full nodes are required. Hence, these are going to be deployed by the partners of the project on more capable hardware. The light smart home nodes and the full partners nodes will altogether make up a fully functional private Ethereum blockchain network. The access to this private network though will be restricted only to GHOST nodes, for security reasons.

Due to the limited number of nodes, the GHOST Ethereum network may be vulnerable to 51% attacks, where a group of mining nodes concentrates more than half of the total hashing power. Apart from the full nodes maintained by partners, only the smart home gateway or middle-ware devices will be allowed to connect to the network. This means that multiple owners of GHOST smart homes must collaborate in order to combine their light nodes to gather more hashing power than what if the servers were combined, in order to theoretically be able to execute a 51% attack. By using enough hardware resources, one can always be sure that the GHOST Ethereum network is secure.

Additionally, in terms of Ether, full nodes will have income from mining while light nodes will have a balance of zero. Ether in the private GHOST Ethereum network do not correspond to real world value. They can be used though to regulate the network usage. Full nodes may equally distribute all or part of their balance to certified GHOST installations nodes, to enable them to commit transactions.

The experimentation network will consist of 10 different nodes; 4 full nodes and 6 light nodes. The full nodes, that act as miners, will have an Intel quad-core 2.2Ghz cpu each, 32GB of memory and 100GB of storage. From the 6 light nodes, 4 will be installed on Raspberry PIs with an assortment of connected Smart Sensors while 2 will be set up on Televes Gateway development hardware. The proposed infrastructure setup is depicted in Figure 2.

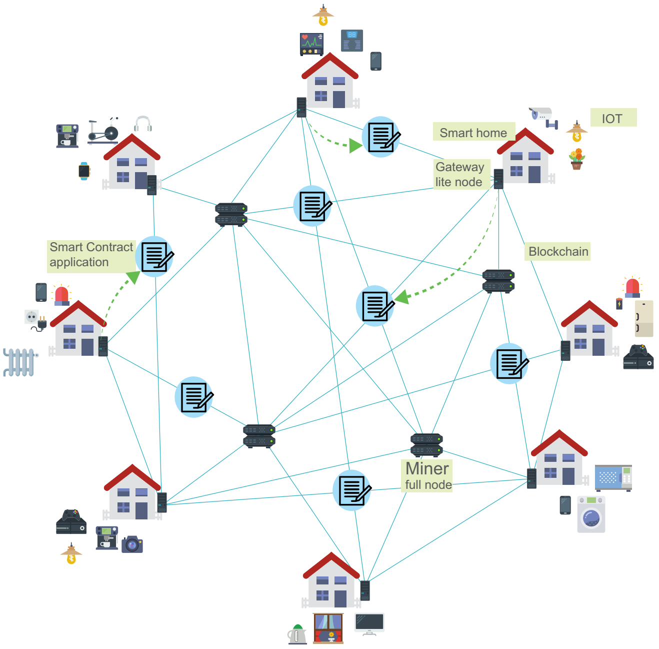

Fig. 2. Infrastructure for SC implementation

A. Blacklisting IPs in a Distributed Sharing

Two types of smart contracts, intended for private and for public blacklisting.

For the Blacklisting IPs use case, there are two relevant entities: CSKB and RE. RE will create blocking rules, based on two input vectors: user and CSKB. An example of a blocking rule is a list of IP addresses maintained and classified as *whitelist* and *blacklist*. The whitelist will contain the IP addresses marked as safe and stored locally on the gateway, further distributed on the SC. The blacklist will contain the IP addresses that are suspected to be malicious. Those among them that come from an online gathered intelligence (public IP addresses) will be added to the SC by CSKB, while the rest will be added to the blockchain by RE through the end-user's configurations and automatic RE decisions. The local copy will be always available on the gateway, and continuous integrity checks will be performed by SC.

1) Public Blacklisting: For the public blacklisting data, a shared and publicly available knowledge of potentially malicious IP addresses will be established. The main storage of the contract will contain a list of records, each one corresponding to the event that a client in a specific installation reports to the GHOST blockchain as malicious. These records cannot be maliciously altered or deleted, as such an action would need to alter data already stored in the blockchain, which is extremely hard. In other words, given that all GHOST nodes act according to the predefined procedure, the information stored in the SC is genuine and may be used from GHOST installations to enhance their risk assessment function.

In terms of the mechanism used to calculate a score (bad reputation) for each external IP, two factors are considered. The first one is the number of existing records related to the specific IP and the second one is the age of such records. It is common for a legitimate host to get infected from malicious software or to be utilised by a remote attacker as intermediary hop in the execution of an attack plan. In these cases the administrator responsible for the host usually discovers the problem and resolves the issue. In such cases, a blacklisting service must be able to remove an IP. In this notion, the SC returns for a specific IP a score (bad reputation), that is calculated according to the number and the age of records for this IP. External hosts that are being reported by multiple GHOST installations will be characterised by high scores. Additionally, when hosts that have been previously reported as malicious, stop appearing in the subsequent records, the corresponding score will eventually decrease and will finally be set equal to zero after a specific time period.

In order for such a mechanism to be functional through an Ethereum SC, specific refinements are required. Too old records, which according to the calculation formula do not have an effect on scores currently produced are of no practical use. These should be discarded, in order to limit the storage resources required for the deployment of the contract and the processing resources required for the execution of the contract. To avoid completely disregarding past information, the discarded records can be stored in a higher level of information. For example a list containing a predefined number of the worst (the ones with highest scores) IPs for each day may be maintained. Such a supplementary layer of information would enable the maintenance of longer history for malicious IPs, without requiring significant resources. In this way it would be possible to additionally penalise old malicious IPs, if they appear again in the future.

Specifically the formula that calculates a bad reputation score for each IP is depicted in Equation 1. By dividing the time in discrete time frames or steps it is easier to implement a scheme that takes into account more recent values with a higher weight. The score is calculated for a specific time period, a specific length of time steps denoted as t_p. If the current time step it t_n, then the score is:

$$score = \frac{\sum_{t=t_n-t_p}^{t=t_n} -\ln{(cf)} sr_t (\lambda)^{t_n-t}}{(\frac{t_p}{trr})^{2r}} \quad (1)$$

The sr_t is equal to 1 if there is a record for the IP in time step t and equal to 0 otherwise. The summation in the nominator does not accumulate values for the time steps at which no record for the specific IP exists.

The λ factor is a decay parameter that takes values in the range (0,1). The higher the value of λ is the strongest the memory of the scheme is. Lower λ values mean that the scoring scheme penalises old values in a more heavy way.

The factor trr in the denominator is the total number of requests that have been made for the specific IP in the time window $[t_n - t_p, t_n]$. The more these requests are, the lower the denominator is, so the score computed from the nominator is increased.

Parameter r denotes if a removal request for the specific IP has recently been submitted. If there has been one, then $r = 1$, otherwise it is equal to zero, $r = 0$. In practice if a removal request has been submitted, then the r factor limits the score. If no recent request exists, then the denominator is equal to 1 and the score is calculated only from the nominator.

Finally, the cf parameter stands for the cardinality factor and penalises the case where all records come from the same submitting address. It is equal to the percentage of records for the specific IP existing in the time window $[t_n - t_p, t_n]$ and have been submitted from the same address as the one being examined. It practically protects the reputation of IPs from spamming accounts, that would want to harm the owner of an IP by repeatedly submitting blacklisting records for this IP.

In order to present the functionality of the formula, some simulation tests have been conducted, the results of which are depicted in Figure 3.

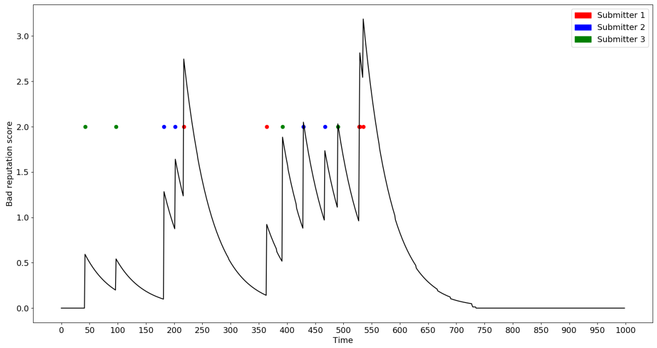

Fig. 3. Reputation score for a specific IP

This figure shows the calculated Bad reputation score for a particular malicious IP, given that the relevant reports for this IP are shown by the coloured dots. The colour of each dot represents the unique id of the submitter of the report. For instance, all red dots are representing submissions of Submitter 1, all blue dots come from Submitter 2 etc.

It is evident in the graph that the scoring scheme values reports according to how recent they are. This is why the score starts to decrease with time after a specific report, at least until a new report is submitted. The rate at which the scheme phases out the past reports is dictated by a decay parameter λ.

Additionally, not all reports contribute the same value to the total reputation score for a particular IP. If a submitter keeps sending reports for the same IP, then every new report is weighted less. This is evident in the case of the first two reports by Submitter 3, the first two green dots, or in the case of the first two submissions by Submitter 2, the first two blue dots approximately at $t = 200$. Consequently, when Submitter 2 keeps quiet for a period of time, his records start to be valued more again, as it is obvious at $t = 470$.

2) Private Black/Whitelisting: The private blacklisting and whitelisting of the IP addresses is a variation of the public blacklisting, where the malicious IP addresses have influence only on a per installation basis. Despite any public recommendation (i.e. Public blacklisting), a user still can have personalised settings and set of rules. Therefore, another set of SC will be put in place for this purpose, where a private list of rules is recorded. Each rule in turn is encrypted together with a state indicating to which list it belongs (i.e. blacklist, whitelist or none for the purpose of resetting the state).

To retrieve the rules, a RE communicates with the SC, which has to aggregate all the rule entries linked with the owner and return a list of the encrypted rules with their entry dates that are stored with the SC. The GHOST client can then decrypt and use the lists to perform several actions, e.g. to restore its internal black and white lists.

B. Implementation Details

There are three main features implemented in the SC for both variations of the blacklisting. The schematic capture of these components is presented in Figure 4, that demonstrates the communication exchange between RE and SC.

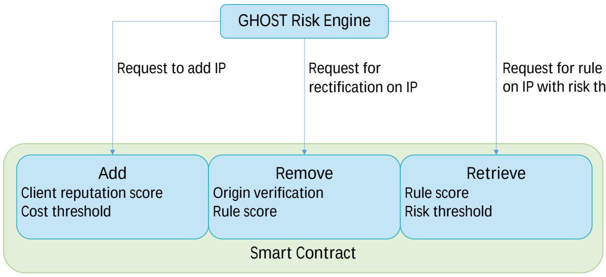

Fig. 4. Smart Contract function outline

1) Reporting Malicious IP: Upon submission of a request to add a new IP to the blacklist, SC will initially calculate the current reputation score of a submitting party, so called *Client*. The SC will also calculate the cost of fulfilling the *Client*'s request. Knowing a consecutive *Client*'s reputation score upon positive outcome, will provide the base for the cost threshold comparison. If the *Client* recently reported the same IP address, there is no need to spend more Ethers on committing the transaction, as well as modifying the related reputation score of the *Client*. Finally, SC will return the outcome of the evaluation (positive or negative) together with a new reputation score of the *Client*, if applicable.

2) Rectifying Erroneous Report: It is possible that *Client* can report the IP address being malicious by mistake as well as by a faulty automated or manual decision. In such a case, a request to remove an IP address from the blacklist can be sent. Such a request should be first verified by the origin, to ensure that only the reporting *Client* can rectify an erroneous action. Once verification is successful, SC will calculate the reputation score of the existing entry, so called *Rule*. If the resulting score exceeds the predefined threshold, a resulting outcome (positive or negative) is sent back to the *Client* together with the new *Rule*'s score if applicable. The threshold is used to avoid situations when committing a new transaction is no longer valuable, as *Rule*'s reputation score won't be significantly improved due to the expiration of the entry's freshness.

3) Gathering Common Dataset: On a regular basis, RE will make requests to the SC to retrieve the *Rule*'s reputation score for IP address of interest. Very often a user's device will initiate communication with an unknown destination. In this case, an acceptable risk level, defined by the GHOST user will be incorporated for deciding on the threshold of the *Rule*'s reputation score. Upon receiving the request on blacklist retrieval, SC will calculate *Rule*'s reputation score and will compare it with the user's defined risk threshold. The result will include only those blacklisted IPs that are below the risk threshold.

The implementation of the smart contracts is outlined in Listing 1 for Public Blacklisting and in Listing 2 for Private Black/Whitelisting respectively.

```
pragma solidity ^0.4.16;

contract PublicBlacklisting {

    struct Record {
        bytes4 ip_address;
        address submitter;
        uint time;
    }

    Record[] public records;

    function submitRecord(bytes4 ip) public {
    .....
    }

    function RemoveRecord(bytes4 ip) public {
    .....
    }

    function evaluateIPScore(bytes4 ip) public {
    .....
    }
}
```

Listing 1. Public Blacklisting smart contract

```
pragma solidity ^0.4.16;

contract PrivateBlackWhitelisting {

    struct Record {
        bytes4 ip_address;
        uint time;
    }

    Record[] public whiteRecords;
    Record[] public blackRecords;

    function submitRecordWhite(bytes4 ip) public
         {
    .....
    }
    function submitRecordBlack(bytes4 ip) public
         {
    .....
    }
    function removeRecordWhite(bytes4 ip) public
         {
    .....
    }
    function removeRecordBlack(bytes4 ip) public
         {
    .....
    }
    function retrieveWhiteList() public view
        returns(Record[]){
    .....
```

Listing 2. Private Black/Whitelisting smart contract

V. Open Issues

The following section focuses on the discussion on the open problems encountered during the design and implementation life-cycle of the Blacklisting IPs use case.

A. Reputation Scoring and Anonymity

The PbD framework employed in the GHOST project dictates specific principles to be applied at all stages of the GHOST solution lifetime. Therefore, during the analysis and design stage, privacy protection measures were considered. First, an anonymous public blacklist submission was considered to prevent GHOST blockchain network neighbours possibility to exploit the full discovery of blacklisted parties. However, this approach directly was leading to an even bigger issue. If one of the network neighbour nodes would happen to have malicious intentions, it could blacklist any node in the network or targeted destination party, compromising the targeted node's communication capabilities for any incoming data from that device. Secondly, the reputation scoring approach currently employed in the SC wouldn't be possible due to the inability to identify malicious reporting party and verify its trustfulness.

B. Spamming on the Blockchain

The nature of the possible spamming within the GHOST blockchain network is similar to malicious blacklisting. One could think of introducing a pre-analysis stage prior to the submission of the blacklist data. Reputation scoring is also providing certain countermeasures to address this problem. However, the introduction of transaction costs for blacklist data submission appears to be most reasonable. In particular, prior to data submission by RE, the associated cost is evaluated based on several criteria: *reporting frequency*, *data freshness* and *common intelligence*. Several strategies are put in place for the dynamic cost calculation, for example making transactions cheaper in case of reporting blacklisting for the party that is publicly known to be malicious (e.g. identified as part of botnet network or being targeted by DDoS attack).

C. Storage on Ethereum

Another important challenge attempted to be addressed is a limitation on the storage capacity of the data on the GHOST private blockchain network. The infrastructure deployment has limitation on the light nodes bound by hardware. This problem is not specific to the GHOST installation only, but is a generic problem in the blockchain community (e.g. technology enthusiasts, developers, business pioneers) and was attempted to be solved by few recent works in [19], [20]. To further evaluate realistic limitations and implement appropriate solution, an in-depth analysis of the possible lifetime is currently performed. In link with privacy profile typology presented in [21], several nodes profiles are considered based on possible activity levels and smart home setup diversity: *fundamentalists*: preserved and sceptical users leaning towards maximum security and minimal exposure, low participation; *unconcerned*: the opposite users, happy to discover new devices and having high trust in technology, high participation; and *pragmatists*: majority of users, continually weighting the benefits between secure exposure and novelty of features provided, medium participation.

D. Costs of Data Submission

As discussed earlier, a need for legitimate strategy on costs and rewards is identified in link with open issues V-B and V-C. One can think of rewarding the nodes on a submission of the blacklist data correlated with publicly available intelligence.

E. Lifetime Availability of the Data and Novel Security

An important generic problem applicable to the whole blockchain community is the potential decryption of the data stored on the blockchain network in the near future. This may happen not only due to technology acceleration and possibility to decrypt data with stronger hardware or quantum computers, but also due to high likelihood of vulnerability discovery in the protocols and software in general. It is therefore possible at some point in the future, for an attacker to discover the black/whitelist data associated to a particular installation and the user. However, by following PbD principles [22], it is always a priority to expose as little personally identifiable data as possible. In particular, even though blacklist reporters can be uniquely identified within GHOST network of installations, the data stored in the blockchain network won't be linking to a unique user globally identified. First an attacker will have to learn the associated unique identifier and real users correlation within the GHOST network, before further possibility of unique tracing of the associated blacklist data.

VI. Conclusions

Given the kind of attention that blockchain technology has received from academia and industry, it appears to be an effective approach to facilitate the future underlying infrastructure for IoT. In this paper, we proposed a smart contracts and blockchain based mechanism to mitigate some of the potential risks associated with the identification and tracking of malicious IP addresses. One of the main contributions of this paper is the use case description and implementation method for two types of smart contracts designed to blacklist malicious IP addresses, publicly and privately. The blacklist will contain the IP addresses suspected to be malicious. For the public blacklist of IP addresses, a shared and publicly available knowledge of potentially malicious IP addresses will be used. On the other hand, for the private blacklist of IP addresses, the list will contain the malicious IP addresses that are likely to have adverse influence only on per installation basis. Despite any public recommendation (i.e. Public blacklisting), a smart home user can personalise the settings and override public rules by whitelisting the party of interest. The second main contribution of this paper is the proposal of a scoring mechanism that serves as the reputation score for the *Rules*. The mechanism calculates a score (bad reputation) for each external IP with the help of two main factors. The first factor is the number of existing records related to the specific IP and the second one is the age of such records. The third main contribution of this paper is the identification of relevant open issues such as reputation scoring and anonymity, spamming on the blockchain, storage on Ethereum, costs of data submission, and lifetime availability of data and novel security. These open issues set the directions for future research, including the validation of the proposed smart contracts by simulation and pilot run.

Acknowledgements

This work is partially funded by the European Union's Horizon 2020 Research and Innovation Programme through GHOST project (https://www.ghost-iot.eu/) under Grant Agreement No. 740923 and by the AAL Programme through Vizier project (https://aalvizier.eu) under Grant Agreement No. AAL-2015-2-145.

References

[1] A. Collen, N. A. Nijdam, J. Augusto-Gonzalez, S. K. Katsikas, K. M. Giannoutakis, G. Spathoulas, E. Gelenbe, K. Votis, D. Tzovaras, N. Ghavami, M. Volkamer, P. Haller, A. Sanchez, and M. Dimas, "GHOST - Safe-Guarding Home IoT Environments with Personalised Real-time Risk Control," in *Proceedings of the 2018 ISCIS Security Workshop, submitted.* Lecture Notes CCIS No. 821, Springer Verlag, 2018.

[2] C. S. Kouzinopoulos, G. Spathoulas, K. M. Giannoutakis, K. Votis, P. Pandey, D. Tzovaras, S. K. Katsikas, A. Collen, and N. A. Nijdam, "Using Blockchains to strengthen the security of Internet of Things," in *Proceedings of the 2018 ISCIS Security Workshop, submitted.* Lecture Notes CCIS No. 821, Springer Verlag, 2018.

[3] P. Johannesson and E. Perjons, *An Introduction to Design Science*. Cham: Springer International Publishing, 2014.

[4] O. Samuel-Ojo, D. Shimabukuro, S. Chatterjee, M. Muthui, T. Babineau, P. Prasertsilp, S. Ewais, and M. M. Young, "Meta-Analysis of Design Science Research Within The IS Community: Trends, Patterns, and Outcomes," *Design Research in Information Systems and Technology*, pp. 124–138, 2010.

[5] W. K. Vijay K. Vaishnavi, *Design science research methods and patterns: innovating information and communication technology*, 2nd ed. CRC Press, 2015.

[6] P. Offermann, S. Blom, M. Schönherr, and U. Bub, "Artifact Types in Information Systems Design Science A Literature Review," in *Global Perspectives on Design Science Research. 5th International Conference, DESRIST 2010, St. Gallen, Switzerland, June 4-5, 2010. Proceedings.*, 2010, pp. 77–92.

[7] A. Whitmore, A. Agarwal, and L. Da Xu, "The Internet of ThingsA survey of topics and trends," *Information Systems Frontiers*, vol. 17, no. 2, pp. 261–274, apr 2015.

[8] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[9] S. Huckle, R. Bhattacharya, M. White, and N. Beloff, "Internet of Things, Blockchain and Shared Economy Applications," *Procedia Computer Science*, vol. 98, pp. 461–466, jan 2016.

[10] B. Lee and J. H. Lee, "Blockchain-based secure firmware update for embedded devices in an Internet of Things environment," *Journal of Supercomputing*, vol. 73, no. 3, pp. 1152–1167, mar 2017.

[11] F. Tian, "A supply chain traceability system for food safety based on HACCP, blockchain & Internet of things," in *14th International Conference on Services Systems and Services Management, ICSSSM 2017 - Proceedings*. IEEE, jun 2017, pp. 1–6.

[12] B. Rodrigues, T. Bocek, A. Lareida, D. Hausheer, S. Rafati, and B. Stiller, "A blockchain-based architecture for collaborative DDoS mitigation with smart contracts," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, ser. Lecture Notes in Computer Science, D. Tuncer, R. Koch, R. Badonnel, and B. Stiller, Eds., vol. 10356 LNCS. Cham: Springer International Publishing, 2017, pp. 16–29.

[13] M. Steichen, S. Hommes, and R. State, "ChainGuard A firewall for blockchain applications using SDN with OpenFlow," in *2017 Principles, Systems and Applications of IP Telecommunications (IPTComm)*. IEEE, sep 2017, pp. 1–8.

[14] J. Tobin, C. Thorpe, and L. Murphy, "An Approach to Mitigate Black Hole Attacks on Vehicular Wireless Networks," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, vol. 2017-June. IEEE, jun 2017, pp. 1–7.

[15] R. Dennis and G. Owen, "Rep on the block: A next generation reputation system based on the blockchain," in *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, dec 2015, pp. 131–138.

[16] Y. Zhao, Y. Li, Q. Mu, B. Yang, and Y. Yu, "Secure Pub-Sub: Blockchain-based Fair Payment with Reputation for Reliable Cyber Physical Systems," *IEEE Access*, pp. 1–1, 2018.

[17] A. Schaub, R. Bazin, O. Hasan, and L. Brunie, "A trustless privacy-preserving reputation system," in *IFIP Advances in Information and Communication Technology*, vol. 471. Springer, Cham, may 2016, pp. 398–411.

[18] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger (EIP-150 revision)(2017)," 2017.

[19] I. Eyal and E. G. Sirer, "Majority is not Enough: Bitcoin Mining is Vulnerable," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9604 LNCS, pp. 106–125, nov 2013.

[20] B. Rodrigues, T. Bocek, and B. Stiller, "Enabling a Cooperative , Multi-domain DDoS Defense by a Blockchain Signaling System," *Semantic Scholar*, 2017.

[21] A. Pedersen, "Usability of authentication in web applicationsa literature review," *July*, pp. 1–33, 2010.

[22] A. Cavoukian and C. Popa, "Embedding Privacy Into What's Next: Privacy by Design for the Internet of Things," *Ryerson University Privacy & Big Data Institute*, no. April, pp. 1–10, 2016.