View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by INRIA a CCSD electronic archive server

# Graph-Based Subjective Matching of Trusted Strings and Blockchain-Based Filtering for Connected Vehicles

Mamoudou Sangaré, Soumya Banerjee, Paul Muhlethaler, Thinh Le Vinh

**To cite this version:**

Mamoudou Sangaré, Soumya Banerjee, Paul Muhlethaler, Thinh Le Vinh. Graph-Based Subjective Matching of Trusted Strings and Blockchain-Based Filtering for Connected Vehicles. MSPN 2020 - 6th International Conference on Mobile, Secure and Programmable Networking, Oct 2020, Paris / Virtual, France. 10.1007/978-3-030-67550-9_1 . hal-02997436

**HAL Id: hal-02997436**

**https://hal.archives-ouvertes.fr/hal-02997436**

Submitted on 10 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Graph-Based Subjective Matching of Trusted Strings and Blockchain-Based Filtering for Connected Vehicles

Mamoudou Sangare [1] Soumya Banerjee [1]
Paul Muhlethaler [1] and Thinh Le Vinh [2]

[1] EVA - Inria Paris. 75012 Paris
emails:{Mamoudou.Sangare,Soumya.Banerjee,Paul.Muhlethaler}@inria.fr

[2] Ho Chi Minh City University of Technology and Education (HCMUTE)
email: hinhlv@hcmute.edu.vn

November 9, 2020

**Abstract**

Advances in technology have led to the creation of a connected world. Due to the increase in the number of smart and autonomous cars and the requirements regarding road safety and associated comfort has led to attempts to adapt conventional vehicular network access to the world of connected vehicles. Consolidating the cooperative safety and collected mobility management from different distributed devices are of the utmost importance. However, the prime objective of connected vehicles is not only to impose security and trust measures for individual vehicles but the strategy of connected vehicles should also concentrate on the cooperative and collective environment of fleets of vehicles. Therefore, keeping simple authentication and access control may not be efficient to evaluate trust and assurance for all the distributed stakeholders. Trust being an important entity for this entire system, the strategy for trust evaluation also becomes crucial. In this paper, we propose a broader content matching model of trusted strings and block chain based filtering for connected vehicles where a content and subject headings are first matched and then the outcome of that is consolidated by a distributed block chain consensus voting mechanism for any decision taken with respect to trust evaluation.

# 1 Introduction

The safety and associated comfort level of driving have motivated the development of connected vehicles. Considering the wide spectrum of connected vehicles, which can communicate in five different modes Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), Vehicle to Cyclist (V2C), Vehicle to Pedestrian (V2P) and Vehicle to Everything (V2X), it is useful to consolidate the cooperative safety and collected mobility management from different distributed devices. However, the prime objective of connected vehicles is not only to impose security and trust measures for individual vehicles, in addition, the strategy of connected vehicles should concentrate on the cooperative and collective environment of a fleet of vehicles. Therefore, keeping simple authentication and access control may not be efficient to evaluate trust and assurance for all the distributed stakeholders. Since trust is an important entity for this entire system, the strategy for trust evaluation also becomes crucial. There are many instances in distributed systems, where trust for multiple parties may not follow the same benchmark for the transmission and reception of messages. This phenomenon could be more prominent, when distributed users carry different mobile edge oriented devices and media. For each of those devices, the transmission and reception strategies with protocols may be different. For example, text messages sent to the mobile devices through social media may not be the same as when sending the same message through mailing or through other types of online media communication. These observations raise some challenges to synchronize distributed mobile edge devices and media against eavesdropping and intentional spam injection procedures. To establish trust for a distributed system, the system should be able to emphasise security aspects and assure the distributed users. The procedure follows a consensus mechanism for the appropriate matching of trusted entities. Dedicated Short Range Communications (DSRC) have been mandatory since 2016 for light vehicles and this rule describes a defined data packet with a Basic Safety Message (BSM) indicating the location of the vehicle, its speed and other on-road parameters. However, DSRC is unable to specify transmitted and received messages with respect to a trusted classification. Therefore, this paper proposes a unique method to investigate the optimal trusted matching for incoming messages in a connected vehicle environment. Interestingly, the paper does not consider key word matching (a word by word or dictionary based approach). Rather, the broader thematic content and headings for communicated messages are taken into account. This will help to establish the content categories for different untrusted behaviors such as abusive behavior, forced branding of products, misleading information, blocking of safety message on road, etc. In order to achieve this matching objective for distributed mobile devices, the paper introduces a message passing procedure followed by a blockchain-based reinforcement decision.

Thus, the paper comprises two parts: the first part describes content-based message passing, and the second part, after matching the content and subject headings, consolidates the distributed consensus or voting mechanism for any decision with respect to the trust evaluation. The key contributions of our paper

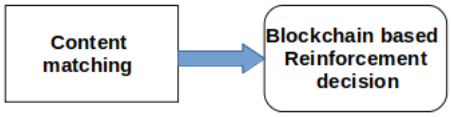


Figure 1: System diagram

are summarized below:

- We propose a message passing scheme for connected vehicles. In this scheme, we do not consider key word matching (a word by word or dictionary-based approach). Rather, we take into account the broader thematic content and headings for messages communicated.
- We attempt to improve trust evaluation by using a voting mechanism for any decision, which is a concept based on a blockchain-based reinforcement decision.
- We aim to enhance securing and authenticating messages exchanged between vehicles by introducing the concept of content matching and trust evaluation in a connected car blockchain as a future perspective.

The remaining part of this paper is structured as follows. The preliminary work done in providing content matching protocols and trust evaluation for connected vehicles is described in Section 2. Section 3 presents the methodology where the entities involved in this work are described. The proposed solution in this study is presented in Section 4, and the experimental results are discussed in Section 5. Finally, Section 6 concludes the paper.

## 2 Related work

Connected vehicle applications are based on both unicast and broadcast communications. However, as for all mobile and wireless networks, these communication scenarios suffer from various security issues that hinder the functionality of such communication protocols. Existing trust-based security solutions are usually classified into entity-based, data-based, and hybrid trust models, depending on the target, which can be dishonest entities, malicious messages, or both of them [6]. In addition, for message passing protocols in vehicular ad hoc networks, especially between connected cars, blockchain technology is seen as the most promising technique to provide secured distributed networks among different frameworks [3]. In the following, we survey message content matching procedures for connected cars as well as providing some background details on blockchain technology for secure message dissemination using a voting mechanism.

## 2.1 Message content matching

In general, string matching has been explored by researchers using different techniques. A technique for detecting phishing attacks was proposed by the authors in [1]. As the objective of that study, this technique was meant to specify the similarity grade between a given URL with blacklisted URLs. Consequently, messages can be classified as phishing or non-phishing based on the textual properties of a URL. In their work, a well-known string matching algorithm called the Longest Common Subsequence (LCS), was implemented by the authors in the hostname for comparison. With an accuracy found to be 99.1%, it is regarded as being very efficient in detecting phishing attacks. It also achieved very low false positive and false negative rates. Similarly, the same algorithm was used in [8]. The authors used it in biological files to discover sequence resemblance between genetic codes. In this test, carried on a sequence of DNA that was generated randomly, the accurate DNA sequence similarity was found by the algorithm. This comparison is a path to implement codes of genetics from one DNA sequence to another. When the algorithm was tested on 50 samples with two input DNA genetic code sequences, it performed well, and showed good results.

The authors in [10] carried out an investigation on the use of string matching algorithms for spam email detection. In particular, their work examined and compared the efficiency of six well-known string matching algorithms, namely Longest Common Subsequence (LCS), Levenshtein Distance (LD), Jaro, Jaro-Winkler, Bi-gram, and term frequency–inverse document frequency (TFIDF) on two various datasets, the Enron corpus and CSDMC2010 spam dataset. From observations based on the performance of each algorithm, they found that the Bi-gram algorithm performed best in spam detection in both datasets. While they claimed that all six methods gave good results in terms of efficiency, however, they suffered from time performance.

The Levenshtein distance algorithm was used by K. Beijering et al. in [4]. They used it to calculate phonetic distances between every 17 Scandinavian language variation and standard Danish. When comparing phonetic transcriptions of two pronunciations, the Levenshtein distance is defined as the number of procedures necessary to convert one transcription to another. The strength of the Levenshtein distance lies in minimising the overall number of string operations when converting one pronunciation to another.

## 2.2 Blockchain technology

A blockchain can be defined as a growing list of records, called blocks, that are linked using cryptography. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data. In other words, it is a distributed and decentralized public database of all transactions or digital events that have been accomplished or shared between participating nodes. Each event

in the public database is validated based on the agreement of a large number of nodes in the blockchain network. The popularity of the blockchain is due to its advantages, which include decentralization, anonymity, chronological order of data, distributed security, transparency and immutability and suitability for trustless environments [9].

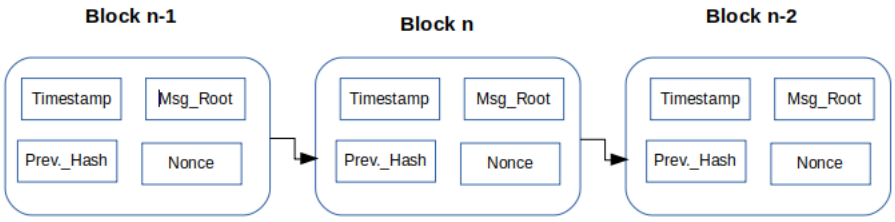


Figure 2: Block chain diagram

The blockchain consists of two types of nodes. A full node is a node that stores and maintains the complete history of blockchain transactions. It begins a transaction directly and independently, and it authoritatively verifies all transactions in the network. Every node in the blockchain network knows the genesis block's hash. Every node in the network builds a trusted blockchain based on the genesis block that acts as a secure root. The genesis block does not have the hash of a previous block. If a node is new, then it only knows the genesis block, and it will have to download all blocks starting from the genesis block to synchronize with the blockchain network and is constantly updated when new blocks are found, see Fig.2. The chaining of blocks is performed by appending hashes of the previous blocks to the current block so that the hash of the current block is in a sequential manner to the following block. Then, it is shared with other nodes in a distributed P2P network in a secure way without the need for a central authority. The sequential hashes of blocks ensure a sequential order of transactions. Therefore, previous transactions cannot be modified without modifying their blocks and all subsequent blocks. The block chain is verified by the consensus of anonymous nodes in the generation of blocks. It is considered secure if the aggregated computational power of malicious nodes is not larger than the computational power of honest nodes. In the case of Bitcoin, the concept of proof of work (PoW) makes sure that a miner is not manipulating the network to make fake blocks. A PoW is a mathematical puzzle that is very hard to solve and easy to verify so that it protects the block chain from double-spending attacks. In the research on VANETs, some of the previous studies related to secure event message dissemination are based on voting. Most voting approaches attempt to solve the issues of node security by asking the opinions of other nodes to determine the trustworthiness of a node.

However, this type of approach has the problem of whether the nodes providing the feedback can be trusted. Generally speaking, limited work has been done to study connected vehicles using the blockchain. The authors in [2] used

a basic blockchain concept to simplify the distributed key management in heterogeneous vehicular networks. The authors in [7] combined the VANET and Ethereum's blockchain-based application concepts and enabled a transparent, self-managed and decentralized system. They used Ethereum's smart contract system to run all types of applications on an Ethereum block chain.

In contrast, our proposed work applies a different type of blockchain for secure message dissemination for connected cars. In [5], the authors proposed a block chain technology for automotive security by using an overlay network in the blockchain and additional nodes called overlay block managers. The overlay network nodes are clustered by cluster heads, and these cluster heads are accountable for handling the block chain and operating its main functions. However, the introduction of additional overlay nodes might cause high latency and might be the center point of failure if the cluster head is compromised.

# 3 Methodology

A comprehensive analysis of message content matching improved by blockchain-based reinforcement decision requires considering multiple entities, e.g. a mobile edge search process that allows us to grab the basic concept of the architecture of mobile edge search process and a graph representation of connected cars. Therefore, this study exploits multiple sources of connected vehicles in terms of message content matching, builds analogous graphs of vehicles' movement patterns for each entity and identifies the community structures.

## 3.1 Architecture of Mobile Edge Search process

The figure below illustrates the architecture of the mobile edge entity search process. Initially the mobile edge entity initiates the handshaking by specifying the sensor observation sequence to be queried by the terminal, and sends the search request to the the mobile edge computing (MEC) server. In return to that request, the cloud server is responsible for responding to the user's search request, and publishing the search request to the MEC server according to the requested content. The MEC server is responsible for fitting the raw data uploaded by the sensor and calculating its similarity with the search conditions published by the cloud server. The sensor layer is responsible for collecting environmental data and uploading it to the MEC server.
Fig. 3 shows the mobile edge entity search process. The steps are as follows:

1. The mobile device reports the environmental message observed to the MEC server.

2. The MEC server fits the reported message of the mobile device, and stores the processed message.

3. The connected car sends a request for an appropriate protocol to the MEC server.

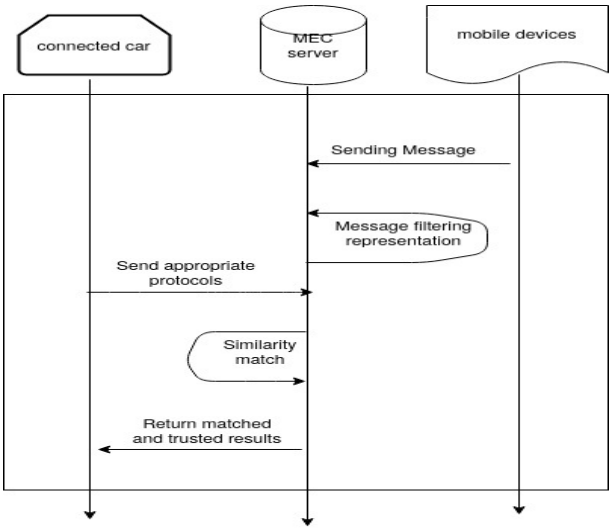


Figure 3: Mobile Edge Entity Search Process.

4. After receiving the request for an appropriate protocol, the MEC server computes the similarity between the search condition and the mobile device message stored internally.

5. Finally, the MEC server returns matched and trusted results that match with the connected car's request to the connected cars.

### 3.2 Graph Representation of Connected Vehicles

A graph is a structure amounting to a set of objects in which some pairs of the objects are in some sense "related". The objects correspond to mathematical abstractions called vertices (also called nodes or points) and each of the related pairs of vertices is called an edge (also called link or line). Fig.4 illustrates a graph representation of vehicles. Nodes(cars) of the graph are in a topological order. For instance in Fig.(4b) we have 1, 4, 6, 5, 2, 3, 7(visual top-to-bottom, left-to-right) or 3, 1, 5, 2, 4 (arbitrary) in Fig.(4a). Each car has an identification number (ID).

## 4 Proposed Solution

The proposed solution of this research will perform trust enhancement among communicating nodes of connected vehicles. The operation comprises two main

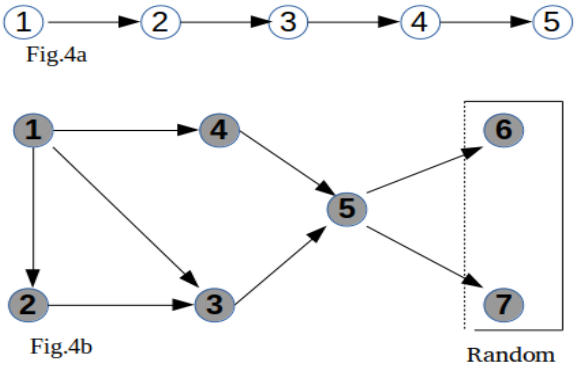


Figure 4: Graph representation of connected cars

Nodes in the graph representation are in topological order.

components, which are content matching under thematic matching operations reinforced by a graph-based blockchain mechanism, see Fig.4.

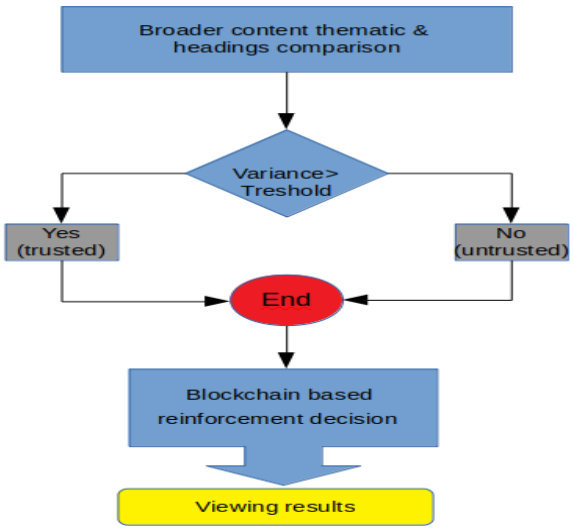


Figure 5: Flowchart of the proposed Solution

The following themes and content are included in the model proposed.
**a) Exhaustive themes** (dangerous product, adult content, gambling and games, inappropriate messaging, personalized promotions, forced promotion)
**b) Non-exhaustive** (affiliating the message against the program rules, promoting the same content from multiple accounts, trying repeatedly to push

Table 1: This table gives a summary of statistical parameters and values.

| | | |
|---|---|---|
| **Data elements** | 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, 4.2, 4.8, 6.4, 6.0 | |
| **mean** | 3.3 | |
| **max** | 6.0 | |
| **variance** | 2.97 | |

brand promotion, brand disinvestment, intentional and manipulation to switch the messages towards inappropriate content).
Under these two heads or leads, the service provider of the connected car can clearly differentiate the two types of content and their thematic message strings. The dictionary is not subjected to one-to-one mapping but it defines lexical matching either in the message head (a) or in the message head (b). This is respective of any theme or content which maybe outside these message heads. This constraint may be a limitation for this model.

## 4.1 Function Matching-Trust

The function of matching-trust is described below:
$(I, S_i, d_I, d_{Si}, d_{min}, \beta)$

---

**Input :**

---

- $I$ is the identifier of the priority string ("xxxx") on the trust graph edge.
- $S_i$ is the string identifier of the moving car transmitting $D_s$
- $d_I$ represents the distance I to the terminating node in case $I \neq$ None (availability steady but trusted).
- $d_{Si}$ represents the distance I to the terminating node in case $S_i \neq$ None (not trusted)
- $d_{min} > 0$ minimum distance of connected cars to perform $D_s$
- $\beta > 1$ co_efficient to transmit the target string

---

**Output:**

---

if($I \neq$ None & $S_i \neq$ None & $d_I > \beta.d_{min}$ & $d_{Si} < d_{min}$) or ($I ==$ None & $S_i \neq$ None & $d_{Si} < d_{min}$) then match_string I
**else**
terminate
**return** Match_string I // untrusted.

---

# 5 Experimentation and Results Analysis

## 5.1 Assumptions

- All connected car members are under the same network service provider on their edge devices.
- Out of the total numbers of members registered in the network, only the agreement of old members (>1 year) could be considered.
- To avoid the physical consensus, the proposed Blockchain prototype will deploy a graph-based referencing. It implies that based on the subjective terms of untrusted message leads, service providers will predefine a maximum high-positive mutual agreement of trusted messages. This typical graph-driven direction will help to prevent latency and delay on the reply of the message block through participants and it also avoids self-biasing to manipulate consensus, if some groups of participants are known to the victim of untrusted acts.

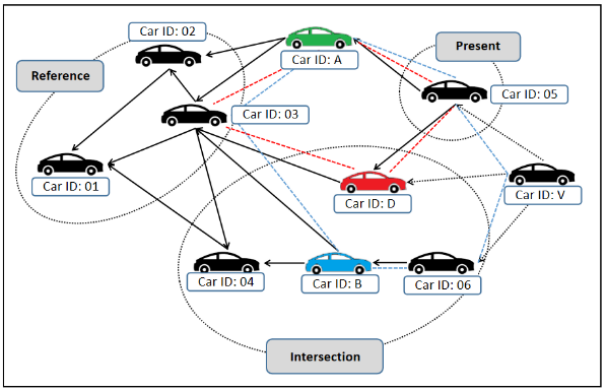


Figure 6: Graph based referencing

Intersections in Fig. 6 conceptually defines that it is the association between certain immediate past values of a car with id transmitted in the message to the neighbors. This includes the values of the car id at present participating in the message transmission. In Fig. 6 the red dotted lines indicate the length of the graph formed either by the transmitting car or its affected neighbors. Therefore they are not trusted. However, distinctly in this cluster, blue dots represent safe messages, where one of the affected cars is placed in the same cluster. In this context, graph referencing is used to investigate the variance and the degree of trust distorted by the odd entry to that cluster.

### 5.1.1 Graph-based referencing towards trusted consensus

The concept here is designed to estimate the temporal inconsistency (ambiguity) between two messages. If two message-clusters are contradictory to each other, their temporal order cannot be determined. This means that the message clusters might be from isolated connected cars. However, the time discrepancy of two message-clusters is bounded by their nearest common ancestor and nearest common descendant, since the real creation time of a message-block is bounded by its ancestors and descendants. The untrusted message-clusters always intend to hide or counterfeit their real creation time in order to carry out spam message generation such as repeat occupancy (conventionally known as double spending). Therefore, the consensus agreement between the untrusted message block and most trusted message-clusters should be very large, otherwise the real creation time of the distrusted block would be bounded by some trusted message-clusters into a small interval. On the other hand, the agreement of two trusted message-clusters is normally much smaller. If the links between message-clusters are not artificially manipulated, the agreement of two message-clusters should only depend on the network propagation speed and the block creation rate. When the network propagation speed or the block creation rate increases, the time discrepancy between the nearest common ancestor and the nearest common descendant will decline. However, the length of shortest path between two message-clusters will increase and cancel out the decline of time discrepancy to some extent. Therefore, the agreements are not very sensitive to the network propagation speed and the block creation rate. The analysis can demonstrate that the agreements between two trusted message-clusters are mostly smaller than 10 while the agreements between the trusted block and the distrusted block might be higher by two or more orders of magnitude. Fig 7 shows the relationship between the block creation rate and the maximum agreement between trusted blocks. In each case of the block creation rate, 16 simulations are conducted. The statistical analysis is shown in Table 1. Even when the message block creation rate reaches 6 message-clusters per second, the agreement between trusted message-clusters still does not increase too greatly. Therefore, the agreements can be utilized to filter the suspect distrusted blocks. In this section, we give a proposed framework named MsgBlock_Filter for identifying the trusted message-clusters based on the agreement. Given a block DAG, we first calculate the agreements for every pair of message-clusters and get the agreement of reference matrix. Then the agreement matrix is converted into a binary matrix where each element is 1 if the corresponding element in the agreement matrix is larger than a preset threshold d, and 0 otherwise. By using the binary matrix obtained as the adjacency matrix, we can construct an undirected graph, in which each vertex represents a block. This graph is called the d-agreement graph of the given block DAG. Intuitively, if a block DAG only contains trusted blocks, the degrees in the d-agreement graph will be very small since the agreements between most trusted message-clusters are zero and the remaining non-zero agreements are also very small. Considering the trusted message-clusters to be the majority, the trusted block identification problem can

be addressed by identifying the maximum subset of vertexes with small degrees. Considering a graph $G = (V, E)$, the $k-$independent set of $G$ refers to the vertex subset $V'$ in which the maximum degree in the induced sub-graph does not exceed k. The maximum $k-$independent set problem is to find the k-independent set with maximum size which is a generalization of classical maximum independent set problem. The maximum k-independent set can be formulated as the following integer programming, in which $xs$ represents whether a certain vertex s is selected and $aij$ denotes the element of adjacency matrix of the graph $G$.

## 5.2 Direct acyclic graph(DAG)

Considering a Direct Acyclic Graph (DAG), it is worth formulating some statistical analysis with respect to message spreading strength (including trusted and untrusted messages) precision and recall. However, due to the legacy of the consensus protocol it becomes more stringent to model the same for different participants in a connected cars environment. The concept for finding the trusted messages and the untrusted or distrusted messages is to find out the interval graph from the first cycle of the message repeat, although the graph here is referred to as an acyclic graph as no cycle exists for the repetition of the message. Therefore the only measure to identify the interval of the message is to find out the variance of the message repeat from one node to another in terms of time. Here we calculate primarily three values for a given message creation rate (Msg_block/second) that is $0.6, 1.2, 1.8, 2.4, 3.0, 3.6, 4.2, 4.8, 5.4, 6.0$ respectively. Under these message creation rates we find the mean to be 3.3, the max to be 6 and the variance to be 2.97. The different steps to calculate the mean, the max and the variance are as follows. Specific points: the variance of any dynamic quantity is the sum of the square difference between each data point and the mean divided by the data value. Hence sigma square should be the sum of the squared difference divided by the total number of items in the given problem. This variance will help to trace the closeness of trusted and untrusted blocks assuming that the untrusted message must be repeated more than once.

- Step 1 : we find the mean of the dataset
- Step 2 : we add all the data values divided by the sample size : $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
- Step 3: we find the sum of the squared difference : $SS = \sum_{i=1}^{n} (x_i - \bar{x})^2$
- Step 4: we calculate variance of sigma squared accordingly : $\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$

In order to identify and grab the described concept of this process, we refer to Fig 7, which shows the maximum number of honest messages versus the message generation rate. Here the variance gives the idea that the density of trusted messages in ideal conditions is always higher. Therefore, even when the message block creation rate reaches 6 message_clusters/second, the variance between trusted messages and clusters become 2.97. Fig 7 also indicates that

the trust level agreement cannot differ too much with respect to untrusted messages. Hence the intersection could be used as a filter for reference to create the predefined trusted and untrusted messages blocks. Two major technical specifications are considered :

1. The predefined referencing of the service provider can prevent the delay in the legitimate reply to the consensus or group messages.

2. Self-biasing or personal manipulation can also avoided.

Fig 8 provides an interesting observation with respect to the precision and the recall by which the strength of the damaging messaging can be highlighted. The left-hand side of Fig 8 is divided almost same intervals apparently. Here also we calculate the quantiles of the given data-set from 0.6 to 6.0 to find out the exact interval of the precision and recall of trusted messages (there is no memory or learning in the recall, only topological ordering has been investigated). Statistically, quantiles are cut points, dividing the range of the data sample of the probability distribution into continuous intervals with equal probabilities. Here in Fig 8 we started calculating the message repeat strength from the median, first quarterly, third quarterly, first decile, last decile, one percentile as maximum level of 6 Msg_Block/second. The flow is to identify the median towards one percentile which is actually the maximum value of the data sample. The analysis helps to correlate the importance of the variance so that repeat messages and the variance can support it as a consensus filter.

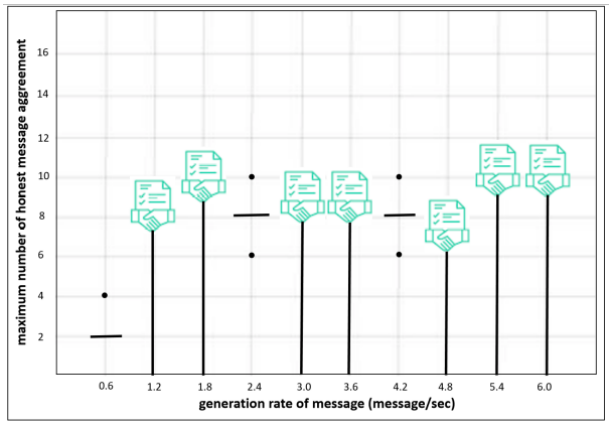


Figure 7: Relationship between the block creation rate and the maximum agreement between trusted blocks

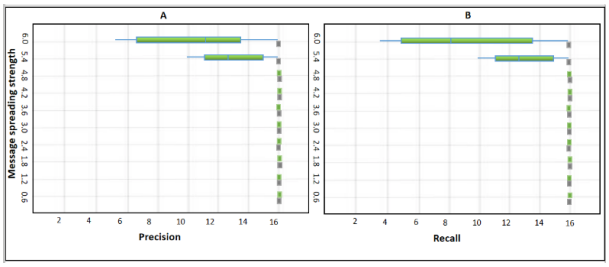


Figure 8: Precision and recall

# 6 Conclusion

In this work, a message matching model and the conceptual level of graph referencing blockchain have been proposed. The model can filter the trusted and untrusted messages in connected car scenarios, analogous to a conventional blockchain mechanism. However, as participants proceed with a voting mechanism, unwanted delay and self-biasing can be introduced in the process. In order to avoid that, a distributed blockchain consensus voting mechanism for any decision taken with respect to trust evaluation is used, this method can be more feasible for collective decisions. This paper has more open research issues challenging the blockchain mechanism. This is because the security is questionable due to group and collective decision-making and repeat occupancy of the message. This is equivalent to a double spending attack in normal blockchain. As a future extension, therefore, a DAG (Direct Acyclic Graph) and the descendants can be integrated in the block-chain, consolidating its security and spoofing mechanism.

# References

[1] Dona Abraham and Nisha S Raj. Approximate string matching algorithm for phishing detection. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2285–2290. IEEE, 2014.

[2] LEI Ao, Chibueze Ogah, Philip Asuquo, Haitham Cruickshank, and SUN Zhili. A secure key management scheme for heterogeneous secure vehicular communication systems. *ZTE Communications*, 14(S0):21–31, 2019.

[3] Muhammd Awais Hassan, Ume Habiba, Usman Ghani, and Muhmmad Shoaib. A secure message-passing framework for inter-vehicular communication using blockchain. *International Journal of Distributed Sensor Networks*, 15(2):1550147719829677, 2019.

[4] Karin Beijering, Charlotte Gooskens, and Wilbert Heeringa. Predicting intelligibility and perceived linguistic distance by means of the levenshtein algorithm. *Linguistics in the Netherlands*, 25(1):13–24, 2008.

[5] Ali Dorri, Marco Steger, Salil S Kanhere, and Raja Jurdak. Blockchain: A distributed solution to automotive security and privacy. *IEEE Communications Magazine*, 55(12):119–125, 2017.

[6] Chaker Abdelaziz Kerrache, Carlos T Calafate, Nasreddine Lagraa, Juan-Carlos Cano, and Pietro Manzoni. Rita: Risk-aware trust-based architecture for collaborative multi-hop vehicular communications. *Security and Communication Networks*, 9(17):4428–4442, 2016.

[7] Benjamin Leiding, Parisa Memarmoshrefi, and Dieter Hogrefe. Self-managed and blockchain-based vehicular ad-hoc networks. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 137–140, 2016.

[8] A Murugan and U Udayakumar. Sequence similarity between genetic codes using improved longest common subsequence algorithm. *International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC)*, 5(7):57–60, 2017.

[9] Rakesh Shrestha, Rojeena Bajracharya, Anish P Shrestha, and Seung Yeob Nam. A new type of blockchain for secure message exchange in vanet. *Digital communications and networks*, 6(2):177–186, 2020.

[10] Cihan Varol and Hezha M Tareq Abdulhadi. Comparision of string matching algorithms on spam email detection. In *2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT)*, pages 6–11. IEEE, 2018.