Evolving Trust Formula to Evaluate Data Trustworthiness in VANETs Using Genetic Programming

Mehmet Aslan[0000−0003−0497−4112] and Sevil Sen[0000−0001−5814−9973]

WISE Lab., Department of Computer Engineering, Hacettepe University, Ankara, Turkey
{mehmetaslan,ssen}@cs.hacettepe.edu.tr

Abstract. Vehicular Ad Hoc Networks (VANETs) provide traffic safety, improve traffic efficiency and present infotainment by sending messages about events on the road. Trust is widely used to distinguish genuine messages from fake ones. However, trust management in VANETs is a challenging area due to their dynamically changing and decentralized topology. In this study, a genetic programming based trust management model for VANETs is proposed to properly evaluate trustworthiness of data about events. A large number of features is introduced in order to take into account VANETs' complex characteristics. Simulations with bogus information attack scenarios show that the proposed trust model considerably increase the security of the network.

Keywords: Evolutionary computation · Genetic programming · Trust management · Data trust · Vehicular ad hoc networks (VANETs)

1 Introduction

Vehicles equipped with smart modules such as Wi-Fi, GPS and computing power can communicate with other vehicles on the road and form a mobile, decentralized, structureless wireless network called Vehicular Ad Hoc Networks (VANETs). Vehicles in VANETs share information to indicate an accident, traffic congestion or another event on a road by sending new event messages or forwarding existing ones to other vehicles. VANET applications use this information to provide traffic safety, improve traffic efficiency and present infotainment.

Complex characteristics of VANETs such as being structureless, decentralized and mobile cause some security challenges. Any vehicle can enter to and exit from a VANET without any control or permission of an authority and start sending or forwarding messages about events occurred on this dynamic environment. This openness makes VANETs vulnerable to several attacks such as bogus information [1]. In this attack scenario, attackers could send false messages about events and also send messages about fake events as if they exist. Vehicles must distinguish bogus information to achieve reliable data transfer within network and maintain traffic safety and efficiency. Trust management model is widely used as a solution against such attacks.

Dynamically changing, decentralized and self-organized topology of ad hoc networks make trust management an optimization problem. Thanks to similar nature of biological systems, nature-inspired optimization algorithms are being used to address some problems in ad hoc networks [2]. A taxonomy of nature-inspired algorithms that used to solve problems in ad hoc networks is given as online/offline techniques, centralized/decentralized systems and proposals using local/global knowledge in [3]. Despite of many aspects of ad hoc networks addressed by evolutionary algorithms, there is a lack of studies that use bio-inspired algorithms to bring a solution to trust management in ad hoc networks.

In this research, a genetic programming based trust management system is proposed to properly evaluate trustworthiness of VANET application data against bogus information attacks. The proposed system uses much more trust evidence than other studies to satisfy the requirements of trust management in VANETs. It selects the most appropriate trust evidences as features to make the right decision about the received event messages and their sender vehicles by evolving a trust calculation formula. The simulation results show that the evolved formula prevents propagation of bogus messages successfully.

The rest of this paper is organized as follows: studies on trust management for VANETs and evolutionary computation (EC) techniques for ad hoc networks are reviewed in Section 2. The network model is explained in Section 3 and the proposed trust management method is described in Section 4. Section 5 gives details of the experiments and presents analysis of the simulation results. Finally, conclusions are drawn and future work direction is presented in Section 6.

2 Related Work

Previous studies are classified into two categories based on their main focus and relevance to this research. Trust management systems for VANETs found in the literature are reviewed in Section 2.1. Section 2.2 highlights evolutionary computation algorithms used to solve different problems in ad hoc networks.

2.1 Trust Management for VANETs

Trust management has many aspects that should be taken into account to establish a proper trust based framework for both VANETs and other ad hoc networks. These aspects are called trust management components and are defined as trust properties, trust management properties, trust metrics and attacks to trust model in several surveys [4,5,6,7,8]. Dynamicity, incomplete/partially transitivity and context-dependency are described in [4,6] and subjectivity and asymmetry are also described in [4] as trust properties. Nonetheless, none of the proposed approaches for VANETs covers all trust properties [4].

In highly dynamic and distributed environments such as VANETs, trust management should be fully decentralized [8]. It is described as one of the most important trust management properties, since a centralized authority cannot be assumed to be exist for trust computation in VANETs [4]. Because of the

possibility of interactions with the same vehicle might be low in a fast and dynamic VANET environment, vehicles cannot wait until direct interactions reach a threshold [8]. Another property that should be considered is capturing dynamicity of VANET environment to calculate the trust based on the current situation using event/task type, location and time information [8]. Moreover, the possibility of uncooperative vehicles to enter VANETs freely should also be taken into account in developing a trust management model [4,8].

Decentralized trust models in VANETs that are based on past interactions and environmental information to take dynamic infrastructure of VANETs into consideration are grouped into three categories: entity-oriented trust models, data-oriented trust models, and hybrid trust models [6,8]. Entity-oriented trust model is the traditional way for trust computing that is proposed for many ad hoc networks including VANETs and MANETs (mobile ad hoc networks). It only considers the trustworthiness of nodes in the network and does not compute different trust values for different messages sent from the same nodes. Calculating only the trustworthiness of messages sent from nodes without considering trust values of the nodes themselves is called data-oriented trust model. Hybrid trust models evaluate both entity and data trust. They use trust value of an entity as a parameter in addition to other trust evidences to evaluate trust values of messages sent from it and also update the entity trust value according to calculated data trust value to maintain a trust relationship based on past interactions.

Wei and Chen [9] propose a hybrid trust model to evaluate the trustworthiness of an event message using beacon trust, event trust and reputation trust values of a vehicle in VANETs. It employs both beacon messages and event messages in order to calculate the trust value, and also update the reputation trust value of vehicles by using the latest event's trust value. Event messages are forwarded either to support or to deny opinion according to a trust threshold in this model. They simulate the model with scenarios including both alteration attacks and bogus information attacks and evaluate the model using F_1 [10] measure. However, they only consider a vector of position, velocity and direction values of vehicle and similarity between event location and estimated location of vehicle as trust evidence with a threshold for distance between receiver and sender and a threshold for time delay between event message time and current time.

Yao et al. [11] proposed an entity-oriented trust model and a data-oriented trust model, however they did not integrate these two models. Even though they use trust value of vehicles in VANETs as a parameter of data-oriented trust model, they do not update the trust value of vehicle using the trust value of data sent from it. They take into account different event types and different vehicle types by assigning weights to them and, introduce a weighted version of successful data forwarding rate using event weight called malicious tendency. This value and vehicle type are then used to calculate trust values of vehicles in the entity-oriented trust model. In order to calculate the data trust, in addition to the trust value of the sender vehicle, they use the distance between the event position and the sender vehicle's position and, the difference between the time of event occurrence and the time of event message. They focus on secure routing in

the simulations of the proposed entity-oriented trust model and, use black hole attack and selective forwarding attack scenarios as well as a network scenario without attacks. Network based metrics are used to evaluate the entity-oriented trust model and an analysis is made for the data-oriented trust model.

To sum up, studies that focus on decentralized trust models in VANETs either take into account very limited trust evidence or do not attach much importance to hybrid trust models. In this paper, we propose a hybrid trust model that mainly aims to evaluate data trustworthiness by using a large number of items of trust evidence that is gathered from the network. Trust values of entities are also calculated based on data trust values of messages sent by these entities.

2.2 Evolutionary Computation Techniques for Ad Hoc Networks

Nature-inspired algorithms developed for solving different problems in ad hoc networks are classified according to their execution mode, information requirement and executing platform in [3]. Firstly, algorithms are classified as online and offline techniques. Secondly, requirement of information about network is considered and algorithms are classified as global knowledge if they need the whole network information and local knowledge if the nodes only use information gathered by themselves. Lastly, optimization algorithms that are run on a central unit are classified as centralized system and optimization algorithms that are run on each node of the network locally are classified as decentralized system. Authors also classified existing studies based on this taxonomy but they did not mention any research about trust management in ad hoc networks. Most of the bio-inspired algorithms used in ad hoc networks are mainly based on two categories, one is centralized and offline with global knowledge and the other is decentralized and online with local knowledge. The latter is more appropriate for trust management in VANETs as each vehicle must evaluate trust values by using only its own local information while moving online on the network.

A recent survey reviews the applications of evolutionary algorithms that is proposed to solve optimization problems in mobile ad hoc networks in the literature [12]. The survey focuses on MANETs, VANETs and DTNs (delay tolerant networks) and divided the reviewed studies into five categories: topology management, broadcasting algorithms, routing protocols, mobility models and data dissemination. Another survey focuses on the applications of evolutionary computation methods for cybersecurity in MANETs [13]. This survey covers evolutionary algorithms (EA), swarm intelligence (SI), artificial immune systems (AIS) and evolutionary games (EG) and classifies these algorithms based on attack types that they counteract and defense mechanisms implemented by them including node trust and reputation systems. It is shown that most of the proposals in the literature is based on EG [13]. The only application of the EA method to trust and reputation systems is proposed for peer-to-peer networks [14]. To sum up, as far as we know the current study is the first application of evolutionary computation techniques to the trust management problem in VANETs.

3 The Network Model

Since there is no well-accepted standard for VANETs yet, an application layer protocol that the proposed trust model is built on is introduced and explained in this section.

3.1 Basic Assumptions

Ad hoc networks are formed by nodes that participate into the network dynamically and contribute to the network communication by behaving both as nodes and as routers. In terms of VANETs, these nodes are vehicles that move at different speeds and generally arrive at different destinations. These nodes encounter other vehicles in the traffic network and make communication with them on the move. Vehicles generally communicate with each other for just a short time, and then never see each other again, which makes the safely communication harder for such dynamic networks. Unfortunately, there is no standard about communication model in VANETs yet, so researchers have been developing new communication models. In the following, some assumptions about vehicles in order to propose a communication model are introduced.

All vehicles have all required devices to communicate with other vehicles over wireless links and form VANET. They could send messages about themselves and events on the road to other vehicles within their communication range. Vehicles have also a unit for calculating trust levels of other vehicles and their messages. Identities and types of all vehicles are also assumed to be controlled and signed by the authorities and this information cannot be changed by vehicles itself.

3.2 Application Model

Many applications running on VANETs mainly focus on sharing information about events that vehicles come across. Vehicles send messages to others while moving on the road to communicate and improve safety and efficiency of the traffic. They mainly send two types of message: beacon and event messages. Events can be considered situations occurred on traffic or road that is worth to share information about them such as a traffic accident, a traffic jam, a toll road or another.

Beacon Messages are periodically sent messages without an observation of an event. Vehicles send beacon messages at every second to their neighbour nodes that are in their direct communication range. This message shows that the sender vehicle of this message is in the traffic network and moving. The beacon message includes current position and velocity data of vehicle at the time of sending this message in addition to unique identifier and type of the vehicle as shown in Table 1.

Table 1. Beacon message format

Vehicle Identifier	Vehicle Type	Message Time	Vehicle Position	Vehicle Velocity

Event Messages are sent by vehicles only when an event is observed. Events that occur in traffic can be categorized into three groups: safety events, efficiency events and infotainment events. Messages about safety events are the most critical type, since it aims to increase traffic safety in critical events such as traffic accident, wet/icy road. Efficiency event messages are sent to establish an efficient traffic network in the case of events such as traffic congestion, road maintenance, closed road. Infotainment event messages carry some information about the facilities nearby such as toll road, scenic area, restaurant, parking/petrol station. Event messages include event type, event description and event position besides the fields exist in beacon messages as shown in Table 2.

Table 2. Event message format

Vehicle Identifier	Vehicle Type	Message Time	Vehicle Position	Vehicle Velocity	Event Type	Event Description	Event Position

3.3 Bogus Information Attack

Suitable security solutions are needed for VANETs in order to overcome vulnerabilities caused by allowing any vehicle to enter to the network such as selfish vehicles, misbehaving ones or even attackers. Selfish vehicles use the network for their own intent. They collect all information from other vehicles but do not send any or send very limited data to them. Their main motivation is using their own resources for only themselves and not being helpful for other vehicles in the network. Misbehaving vehicles could have some malfunctioned device or could be captured by an attacker and send false information unintentionally. Vehicles that aim to damage the network deliberately are called attackers.

In this study, attackers carry out bogus information attacks in order to harm the network. In this attack scenario, even though the attackers observe events like other vehicles, they do not send genuine messages about the events they encounter with. Instead, they send fake and false information about an existent or nonexistent event to their neighbours. Attackers modify the event type of a real event in order to mislead their neighbours. They also generate and send fake event messages with event type, event description and event position data in order to gain some advantage on the road. For example, they could decrease the density of the road they have been using by sending fake messages about a nonexistent accident on that road. Vehicles should be aware of that kind of attackers and they must decide whether the received messages from such nodes are trustable or not. Proposing a trust management model against such attacks is the main motivation of this study.

4 The Proposed Method

Trust management models are widely used by researchers in ad hoc networks in order to ensure secure and reliable communication. In such models, each node assigns a trust degree to the messages they receive and/or the nodes that the message is received from. Trust formula is used to calculate such trust degrees by using the available information in the network. However, generally manually generated trust formulas will have limited number of features and, hence cover only a little aspect of network. They will not be able to represent complex properties of VANETs. A trust management model proposed for VANETs should be able to reflect changes in topology and events in the model.

In this study, we investigate the use of genetic programming in order to generate a trust management model automatically in order to efficiently and effectively handle dynamically changing topology and events of VANETs. This trust formula is generated using more features than other studies in the literature. The features are selected to represent complex characteristics of this dynamic environment. The components of the proposed trust management model are described in the following sections.

4.1 Vehicle Type and Weight

Vehicles in VANETs have different roles and objectives on traffic based on their types. They are divided into three groups: police automobiles, public service vehicles and ordinary vehicles. Vehicle type usually indicates trustworthiness of vehicles to some extent. Police automobiles are responsible of controlling the traffic and providing road safety, therefore they are the most trustworthy vehicles in the network. Public service vehicles such as ambulance, bus, engineering vehicle, etc., are usually on duty for ensuring either road safety or efficiency. They are considered as medium level vehicles in the proposed trust model. Ordinary vehicles such as private cars, taxis, etc., are considered as low level vehicles from the trust point of view, since their contribution to road safety is generally lower than others. In order to use this knowledge on trust calculations, a trust feature called vehicle weight $W_V(x)$ is defined as in Eq. 1:

$$W_V(x) = \begin{cases} 1.0, & \textit{when } x \textit{ is a police automobile} \\ 0.7, & \textit{when } x \textit{ is a public service vehicle} \\ 0.5, & \textit{when } x \textit{ is an ordinary vehicle} \end{cases} \tag{1}$$

4.2 Event Type and Weight

Events have different impacts on traffic and road safety and require different trustworthiness levels. The most important message type is clearly safety events as described above. Vehicles in VANETs pay attention to messages' importance levels in order to maintain road safety. This information is represented with a

trust feature called event weight $W_E(x)$ as defined in Eq. 2:

$$W_E(x) = \begin{cases} 1.0, & when\ x\ is\ a\ safety\ event \\ 0.8, & when\ x\ is\ an\ efficiency\ event \\ 0.5, & when\ x\ is\ an\ infotainment\ event \end{cases} \quad (2)$$

4.3 Trust Evidence

Each term in the trust formula expression is called trust evidence. They represent the characteristics of network including the properties of vehicles and messages. Each vehicle participated into VANET gathers items of evidence about network by using both beacon and event messages. The values of items of trust evidence used in this study are normalized to $[0, 1]$. The trust formula is based on such trust evidence calculated by using messages received from the neighbour nodes. In order to prevent unnecessary computing overhead, the calculation of the trust value takes place only when a vehicle receives an event message. In addition, beacon messages are stored in a sliding window of 5 messages and stale messages are discarded to keep the memory consumption low. Table 3 shows the trust evidence set and Table 4 lists the notations used in the model.

Vehicles calculate the neighbourhood density as the ratio of the number of current neighbours to the encountered maximum number of neighbours by that time. The percentage of newly added neighbours and removed neighbours since the delivery of the last event message is also monitored.

Position and time proximities are important factors to decide whether the trust value of an event message or its sender should be calculated or not. Some messages are not taken into account for the calculation of trust value according to their position and time proximity.

An event could be observed through many messages sent from more than one vehicle. When an event message is received, the receivers wait for a fixed time to receive other messages of the same event. Since these messages are valuable to calculate the trustworthiness of the received messages, there are also some items of trust evidence based on them as shown in Table 3.

4.4 Trust Distribution

The dynamically changing topology of VANETs could cause vehicles to encounter with vehicles that they have not communicated before and had no prior experience about. Therefore, they should prefer to take in consideration the recommendations from their own trustee rather than deciding randomly to trust such newly encountered vehicles or not. Trust distribution plays a vital role to achieve that. Vehicles only forward a message that they have decided to be trustworthy. Before they forward the message, they add their opinions about the message and its sender. This opinion contains both the trust value of the event and the trust value of its sender. Besides these two trust values, the following information about the forwarder node is also added to the event message: its identifier, type, position and velocity. Table 5 shows the forwarded event message format.

Table 3. Trust evidence set

Notation	Trust Evidence
ND	Neighbourhood density
ANP	Percentage of added neighbours
RNP	Percentage of removed neighbours
EP	Proximity of the receiver to the event
VP	Proximity of the receiver to the source vehicle
TP	Event time proximity
W_V	Weight of the source vehicle
W_E	Weight of the event
ET	Recommendation (trust value) of the event sent by the forwarder
SP	Percentage of the nodes sending the same event
SW	Average weight of the nodes sending the same event
EW	Average weight of the events at the same location
TE	Average weighted forwarder recommendation about event trust
TV	Average weighted trust value of the sender vehicle

Table 4. Notations

Notation	Definition
NN_A	number of neighbours of vehicle A
MN_A	maximum number of neighbours of vehicle A
$ND_A = NN_A \ / \ MN_A$	neighbourhood density of vehicle A
AN_A	added number of neighbours of vehicle A
RN_A	removed number of neighbours of vehicle A
$ANP_A = AN_A \ / \ MN_A$	percentage of added neighbours of vehicle A
$RNP_A = RN_A \ / \ MN_A$	percentage of removed neighbours of vehicle A
ED_A^X	distance of vehicle A to the event X
VD_A^B	distance of vehicle A to the source vehicle B
MD	maximum allowed distance
$EP_A^X = (MD - ED_A^X) \ / \ MD$	proximity of vehicle A to the event X
$VP_A^B = (MD - VD_A^B) \ / \ MD$	proximity of vehicle A to the source vehicle B
T	current time
GT_X	generation time of the event message X
MT	maximum allowed event time
$TP_X = (MT - (T - GT_X)) \ / \ MT$	proximity of event X to current time
W_V^A	weight of the source vehicle A
W_E^X	weight of the event X
VT_A^B	trust value of vehicle B calculated by vehicle A
ET_A^X	trust value of event X calculated by vehicle A
SN_A^X	the number of nodes sending the same event X
$SP_A^X = SN_A^X \ / \ MN_A$	percentage of nodes sending the same event X
$SW_A^X = (\sum_{i=1}^{SN_A^X} W_V^i)/SN_A^X$	average weight of nodes sending the same event
$EW_A^X = (\sum_{i=1}^{SN_A^X} W_E^i)/SN_A^X$	average weight of events at the same location
$TE_A^X = (\sum_{i=1}^{SN_A^X} VT_A^i * ET_i^X)/SN_A^X$	average weighted event trust value
$TV_A^X = (\sum_{i=1}^{SN_A^X} VT_A^i * VT_i^B)/SN_A^X$	average weighted sender trust value

Table 5. Forwarded event message format

Vehicle Identifier	Vehicle Type	Message Time	Vehicle Position	Vehicle Velocity	Event Type	Event Description	Event Position
Forwarder Identifier	Forwarder Type	Forwarder Position	Forwarder Velocity	Forwarder Event Trust		Forwarder Sender Trust	

When vehicles receive a forwarded event message, they calculate the average trust value of the sender vehicle weighted by the trust values sent by forwarders. The average trust value of the event is also calculated based on the forwarders' opinions about it. All direct and recommended trust evidences are used to compute the combined trust value. These recommended trust evidences are also shown in Table 3.

4.5 Trust Update

Vehicles assign trust values not only to event messages but also to the senders of those messages. Messages sent from vehicles that have higher trust value are decided more likely to be trustworthy than messages from untrusted vehicles. At the beginning, the trust value of each vehicle is set to 0.6. Every time a message is received from the sender, its trust value is updated according to the Eq. 3 given below. Interactions with the sender is taken into account for trust update. Let's assume that the vehicle A receives a message sent from the vehicle B. Here, (ET_A^B) represents the trust value of this message and TT refers to the threshold for accepting this message to be forwarded. Where (CT_A^B) shows the current trust value of the sender vehicle B, (NT_A^B) indicates the newly updated trust value of the sender vehicle B calculated by the receiver vehicle A.

$$NT_A^B = \begin{cases} CT_A^B \times \dfrac{ET_A^B}{TT}, & 0 \leq ET_A^B < TT \\ CT_A^B + (ET_A^B - CT_A^B) \times (\dfrac{ET_A^B - TT}{1 - TT}), & TT \leq CT_A^B \leq ET_A^B \leq 1 \\ CT_A^B, & TT \leq ET_A^B < CT_A^B \leq 1 \end{cases} \tag{3}$$

Well-known principle about trust "hard to earn but easy to lose" [6,7,8] is applied while calculating (NT_A^B). Vehicles must send messages that are more trustworthy than vehicles itself to increase their trust values. Even a message is considered as trustworthy; it will not change the trust value of its sender unless its trust value is higher than vehicle's. Increasing rate of vehicle's trust value is proportional to gap between event trust value and vehicle trust value, and the normalized trust value of the message. In contrast, untrusted messages will decrease the trust value of sender rapidly.

4.6 Evolving Trust Formula by using Genetic Programming

Genetic programming (GP) [15,16] is a population-based search algorithm inspired by natural evolution. It starts with generating a population of individuals (usually at random) which are candidate solutions for the target problem. Then, each individual is evaluated and assigned with a fitness value that indicates how well this candidate solves or comes close to solving the problem at hand. Until a termination criterion is satisfied, new populations are generated iteratively by using selection, crossover, and mutation operators, as in natural evolution. These genetic operators are used to provide better solutions in the new population.

Each individual, candidate solution for the problem in other words, represents a trust formula, which is generated randomly at first generation. Each individual is represented as a tree in GP. In-order traversal of the tree outputs a candidate trust formula. Terminal nodes are trust evidences in Table 3 and some ephemeral random constants (ERC). Non-terminal nodes consist of mathematical operations listed in Table 6. These operations are implemented to have the result value of $[0, 1]$. Each individual is assigned a fitness value based on its detection rate of false and fake messages. Higher value of fitness value shows better individuals, so the algorithm tries to increase the fitness value of population using genetic operators. Selection operator probabilistically determines the parent individuals that will be used in the crossover and mutation operators. Better individuals have a higher chance to be selected. Crossover and mutation operators are used on the selected parents to breed new individuals. The crossover operator exchanges different portions of the parents and produces two new child individuals. It aims to create better solutions using good parts of parents. In the mutation operator, some portions of newly generated solutions are changed randomly in order to increase diversity and reach better solutions.

Table 6. Genetic programming operation set

Add $(X + Y) / 2$	Mult $X \times Y$	Square $X \times X$	Cube $X \times X \times X$	Neg $1 - X$
Sub $(X - Y + 1) / 2$		Exp $(e^X - 1) / (e - 1)$		Sqrt $\sqrt{X}$
Sin $(\sin(\pi X - (\pi / 2)) + 1) / 2$			Cos $(\cos(\pi X) + 1) / 2$	

5 Experiments

The experiment consists of two parts: network simulation and the evolution of the trust formula by using genetic programming. The mobility of vehicles on the road, data transfer between vehicles, trust computation of vehicles about other vehicles in neighbourhood and bogus information attack scenario are simulated using the ns-3 network simulator [17]. The ECJ toolkit [18] is used for genetic programming, which automatically generates trust formula as candidate solutions and computes the fitness values of such solutions after running the

each evolved formula in the network simulation. By using the evolved trust formula, vehicles determine whether messages are trustworthy or not in a network scenario where both normal and fake messages exist.

5.1 Network Simulation

In the simulation, vehicles are generated and moving according to a real world traffic model taken from a street map in Zurich. This real world traffic model [19] is included in the distribution of ns-3. Each simulation has one of the traffic density settings as low, medium and high and takes 300 seconds.

Vehicles send two types of messages to others: beacon and event messages. Beacon messages are sent periodically at every second. Event messages are sent when an event occurs in the 100m range of the node. Beacon messages are stored in a sliding window of 5 messages and processed when an event message is received. Vehicles process these messages and obtain the values of trust evidences listed in Table 3 in order to calculate both the trust value of the sender and the trust value of the message. Trust threshold is set to 0.6 and the event message is forwarded if the vehicle trusts the message. The forwarding vehicle inserts its own trust value about both the sending vehicle and the message. Table 7 shows the parameters of the network simulation.

Table 7. Network simulation parameters

Name	Value
Simulation area	4.6 km x 3.0 km street map
Number of vehicles	99 (low), 210 (medium), 370 (high)
Vehicle types	high, medium, normal
Ratio of high level vehicles	5%
Ratio of medium level vehicles	15%
Ratio of attackers	10%
Attack type	bogus information
Simulation time	300 seconds
Mobility	real traffic data model
Beacon interval	1 second
Beacon window's size	5 messages
Beacon messages' size	128 bytes
Event messages' size	256 bytes
Event types	safety, efficiency, infotainment
Ratio of safety events	10%
Ratio of efficiency events	40%
Event detection range	100 meters
Max event distance	500 meters
Max event time	1 seconds
Max delay time	0.2 seconds
Routing protocol	none
Trust threshold	0.6

5.2 Evolution of Trust Formula

Each individual is run on the network simulation in order to calculate its fitness value. By using the evolved trust formula, vehicles determine whether messages are trustworthy or not in a network scenario where both normal and fake messages exist. A vehicle makes true positive (TP) decision if it decides a malicious event message is untrustworthy. In contrast, if vehicle decides a normal event message is trustworthy, it makes true negative (TN) decision. A vehicle makes false positive (FP) decision if it decides a normal event message is untrustworthy. On the other hand, if vehicle decides a malicious event message is trustworthy, it makes false negative (FN) decision. After obtaining TP, TN, FP and FN values, precision rate and recall rate are calculated and fitness value of a generated trust formula is determined using F-measure (F) [10], defined as in Eq. 4. It takes value in the interval $[0, 1]$. Table 8 shows the parameters of genetic programming. The parameters not listed here are the default parameters of the ECJ toolkit.

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4}$$

Table 8. Genetic programming parameters

Parameter name	Parameter value
Population size	100 individuals
Maximum generation number	20
Crossover probability	0.7
Mutation probability	0.3
Tournament size	7
Terminal nodes	Trust evidences and ERC
Nonterminal nodes	add, sub, mult, sin, cos, exp, square, sqrt, cube, neg

5.3 Results and Analysis

The GP algorithm is run five times. In each run, different network settings having different events are applied. The best individual of these runs of GP is evaluated here. The change in the fitness value of the best individuals of different runs is shown in Fig. 1. The best trust formula, which is evolved by using a traffic network under low density (99 vehicles), is given in Table 9. As shown in the formula, 8 of 14 trust evidences, 8 of 10 operations and an ERC have selected in the evolution process.

The best trust formula is evaluated on networks with different event positions and varying density patterns from low density of 99 vehicles to high density of 370 vehicles as shown in Fig. 2 and Fig. 3 respectively. Even though individuals are trained in a low density network, the best performances of the formula is achieved on networks under high density. It is an expected result, since a node has a higher

Table 9. The evolved best trust formula

Mult(Neg(Sqrt(averageWeightedSenderTrustValue)), Add(0.7483884231631781, Mult(Sqrt(Add(Add(neighbourhoodDensity, Sin(receiverToSenderProximity)), Mult(Sqrt(Sqrt(Add(averageEventTypeWeight,Mult(Add(receiverToEventProximity, averageWeightedSenderTrustValue), Add(Neg(Sub(Cos(Add(eventTimeProximity, senderTypeWeight)), Mult(Exp(averageSenderTypeWeight), neighbourhoodDensity))), Exp(averageSenderTypeWeight)))))), Sin(receiverToSenderProximity)))), Sin(receiverToSenderProximity))))

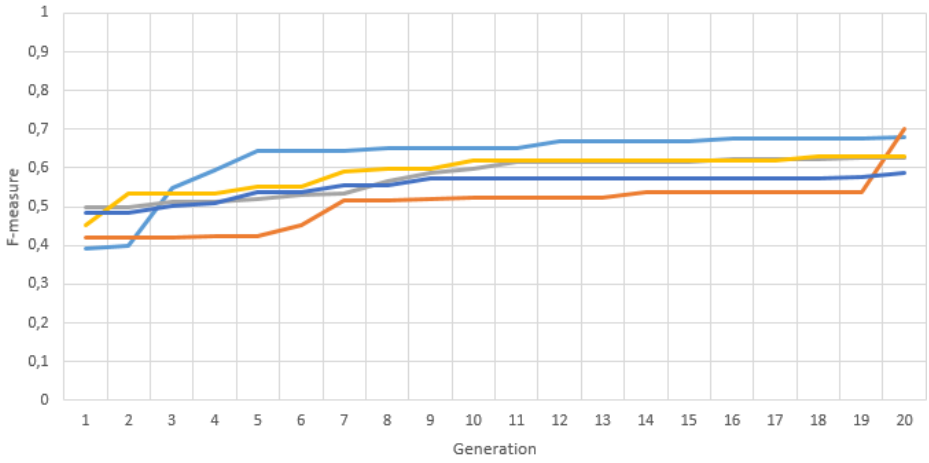

Fig. 1. Change in the fitness value of best individuals over generations

number of neighbours on dense networks; they get more recommendations about a node or a message.

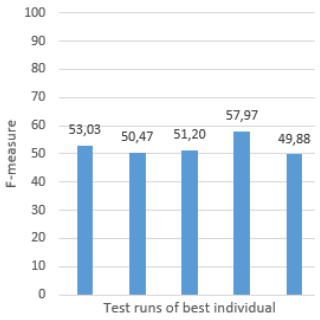

Fig. 2. Performance under low density

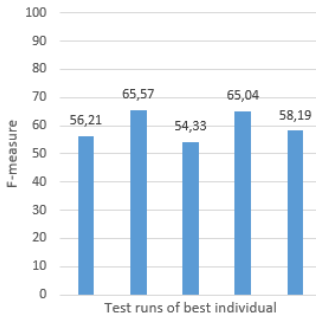

Fig. 3. Performance under high density

6 Conclusion

This paper presents the first study that explores the use of evolutionary computation techniques to the trust management problem in VANETs. A method based on genetic programming is proposed in order to evaluate data trustworthiness of events in VANETs automatically. A large number of features collected from both event messages and beacon messages in the network are introduced in order to discover complex properties of VANETs. The feature set covers much more trust evidence than other studies in the literature. A trust formula based on this feature set is evolved by using genetic programming. The simulation results shows that the proposed model is effective against bogus information attacks.

References

1. Sakiz, F., Sen, S.: A survey of attacks and detection mechanisms on intelligent transportation systems: VANETs and IoV. Ad Hoc Networks 61, 33–50 (2017)
2. Dorronsoro, B., Ruiz, P., Danoy, G., Pigné, Y., Bouvry, P.: Evolutionary algorithms for mobile ad hoc networks. John Wiley & Sons (2014)
3. Dorronsoro, B., Ruiz, P., Danoy, G., Pigné, Y., Bouvry, P.: Survey on Optimization Problems for Mobile Ad Hoc Networks, chap. 3, pp. 49–78. John Wiley & Sons (2014)
4. Cho, J.H., Swami, A., Chen, R.: A survey on trust management for mobile ad hoc networks. IEEE Communications Surveys & Tutorials 13(4), 562–583 (2011)
5. Govindan, K., Mohapatra, P.: Trust computations and trust dynamics in mobile adhoc networks: A survey. IEEE Communications Surveys & Tutorials 14(2), 279–298 (2012)

6. Ma, S., Wolfson, O., Lin, J.: A survey on trust management for intelligent transportation system. In: Proceedings of the 4th ACM SIGSPATIAL International Workshop on Computational Transportation Science. pp. 18–23. ACM (2011)
7. Yu, H., Shen, Z., Miao, C., Leung, C., Niyato, D.: A survey of trust and reputation management systems in wireless communications. Proceedings of the IEEE 98(10), 1755–1772 (2010)
8. Zhang, J.: A survey on trust management for VANETs. In: Advanced information networking and applications (AINA), 2011 IEEE international conference on. pp. 105–112. IEEE (2011)
9. Wei, Y.C., Chen, Y.M.: Efficient self-organized trust management in location privacy enhanced VANETs. In: International Workshop on Information Security Applications. pp. 328–344. Springer (2012)
10. Van Rijsbergen, C.J.: Information Retrieval. Butterworths, London, UK (1979), http://www.dcs.gla.ac.uk/Keith/Preface.html, Last accessed 31 Jan 2019
11. Yao, X., Zhang, X., Ning, H., Li, P.: Using trust model to ensure reliable data acquisition in VANETs. Ad Hoc Networks 55, 107–118 (2017)
12. Reina, D.G., Ruiz, P., Ciobanu, R., Toral, S., Dorronsoro, B., Dobre, C.: A survey on the application of evolutionary algorithms for mobile multihop ad hoc network optimization problems. International Journal of Distributed Sensor Networks 12(2), 2082496 (2016)
13. Kusyk, J., Uyar, M.U., Sahin, C.S.: Survey on evolutionary computation methods for cybersecurity of mobile ad hoc networks. Evolutionary Intelligence 10(3), 95–117 (2018)
14. Tahta, U.E., Sen, S., Can, A.B.: Gentrust: A genetic trust management model for peer-to-peer systems. Applied Soft Computing 34, 693–704 (2015)
15. Koza, J.R.: Genetic programming: On the programming of computers by means of natural selection. MIT Press, Cambridge, MA (1992)
16. Koza, J.R.: Genetic programming as a means for programming computers by natural selection. Statistics and computing 4(2), 87–112 (1994)
17. The NS-3 Network Simulator, https://www.nsnam.org/, Last accessed 31 Jan 2019
18. Luke, S.: ECJ A Java-based Evolutionary Computation Library (1998), https://cs.gmu.edu/~eclab/projects/ecj/, Last accessed 31 Jan 2019
19. Naumov, V., Baumann, R., Gross, T.: An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces. In: Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing. pp. 108–119. ACM (2006)