Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 14, No. 3, 2025

Spatio-Temporal Graph-Based Association Rule Mining Method for Anomaly Node in Internet of Vehicle

Yangmei Zhang *, Zixuan Wang

Solution Research Center, China Post Era Telecom Technology Co., Ltd. Beijing, China
*** Corresponding author:** Yangmei Zhang (Email: 1203265381@qq.com)

Abstract: Abnormal behaviors in the Internet of Vehicles often manifest as sudden changes in communication behavior, abnormal fluctuations in communication volume, and other irregular variations. These anomalies may be signs of a cyberattack, such as attackers attempting to disrupt normal commu-nication between vehicles by severing connections, establishing unauthorized communication links, or sending a massive amount of data packets to exhaust resources or isolate specific nodes. Traditional association rule methods struggle to handle the complex spatio-temporal characteristics, large-scale datasets, and the dynamic communication behavior among vehicle nodes that change over temporal and spatial. This study introduces spatio-temporal features to capture the communication behavior of vehicles as they vary across temporal and spatial, proposes an association rule mining algorithm based on spatio-temporal graphs, which deeply combines the spatio-temporal features of the Internet of Vehicles, such as location, time, communication behavior, and communication volume, to precisely track and analyze changes in vehicle node communication with the goal of identifying anomalous behavior nodes. Utilizing spatio-temporal graph-based association rule mining technology reveals nodes with anomalous behaviors characterized by frequent communication and dramatic increases in communication volume. The algorithm is designed in two steps: discovering frequent rules and verifying the credibility of these rules. Thanks to the adoption of connection and elimination strategies, this algorithm demonstrates the efficiency and accuracy in rule mining sought after in the field. Compared to traditional methods, our proposed mining algorithm shows significant improvements in mining time and the accuracy of the number of rules mined, especially displaying a remarkable advantage in handling the complex spatio-temporal characteristics of the Internet of Vehicles.

Keywords: Internet of Vehicles; Spatio-Temporal Features; Association Rule; Communication Behavior; Anomaly Node.

1. Introduction

Anomalous behaviors in the Internet of Vehicles (IoV) [1] can be considered from the perspective of network communication abnormalities, which are mainly reflected in sudden changes in communication behavior, abnormal increases and decreases in communication volume, or abnormal fluctuations in communication content. Sudden changes in communication behavior may indicate that attackers are attempting to disrupt the normal communication flow between vehicles to carry out further attacks or interference. For example, attackers might sever the disconnect a vehicle from the network or establish an illegal communication link between vehicles. Abnormal increases or decreases in traffic are often a sign that network resources are being maliciously consumed or intercepted. Attackers may deplete the processing capabilities of targeted vehicles or infrastructure by sending a large communication volume of meaningless data packets, or they might reduce the communication volume arriving at specific nodes, thereby isolating the target node. Abnormal fluctuations in communication content may imply data tampering or the risk of information leakage [2].

Attackers may modify data packets in transmission to send false information or steal sensitive data. These anomalous behaviors are primarily due to network attack methods, making the accurate identification and timely response to these anomalous vehicle nodes crucial.

Spatio-temporal graphs, as a variant of graph data structures, has become a key tool for understanding and analyzing the dynamics of the Internet of Vehicles system by introducing the dimensions of temporal and spatial to capture the changes of nodes and edges over time and location. In the IoV context, temporal characteristics reflect changes in vehicle behavior over time. For instance, by analyzing the communication behavior of vehicles (referring to the frequency of information transmission and communication volume) at different times, one can identify anomalous communication behaviors among nodes. Spatial features [3] focus on the movement of vehicles in location space. By analyzing the spatial relationship between vehicles, the coverage and connection stability of the network can be understood. Therefore, the dynamic characteristics [4] of the IoV are comprehensively analyzed by combining temporal and spatial characteristics. For example, spatio-temporal Features can help identify the network communication behavior among vehicles at specific locations and times. Nodes (vehicles) and edges (communications between vehicles) in a spatio-temporal graph [5, 6] are assigned timestamps and geographical locations, enabling the representation of how vehicles move over temporal and spatial. For example, considering the communications among vehicles A, B, and C, a dynamic spatio-temporal [7] graph could include the following information: 1) At time point t1, vehicle A is at geographical location L1, sending a message volume Q1 to vehicle B; 2) At time point t2, vehicle B is at geographical location L2, communicating with vehicle C, assuming the message volume is Q2. This dynamic spatio-temporal graph not only records the geographical location and time points of the vehicle, but also includes the dimensions of communication behavior and communication volume, which can be used to analyze dynamic changes in inter-vehicle communication, such as frequent communications, changes in communication volume, etc. Through this changing trend is used to dig out abnormal nodes.

Sudden changes in communication frequency may suggest frequent disconnections and reconnections of a specific vehicle node, primarily caused by Denial of Service (DoS) attacks [8] or failures [9] in network equipment. Without apparent demand changes, a significant increase in communication volume might indicate signs of a flooding attack. Normal IoV communication content should reflect the real-time status and control information of vehicles. If communication content shows illogical fluctuations, such as abrupt changes in location information, it may signal that vehicle nodes have been tampered with or information has been maliciously altered. By constructing dynamic spatio-temporal graphs, these abnormal behaviors can be analyzed more intricately. Dynamic spatio-temporal graphs enable the comprehensive consideration of a vehicle's location, time, communication behavior, and communication volume, thus providing a multidimensional perspective for observing and analyzing anomalous behaviors in the IoV.

Temporal association rule mining[10] is a complex task, especially when the spatio-temporal attribute [11] characteristics of vertices in a graph sequence need to be considered. At present, there have been studies to solve this problem by defining temporal association rules of graphs and proposing corresponding mining methods, but most of these studies focus on single-attribute graph sequences. Although these methods are effective to a certain extent, they still face challenges in handling spatio-temporal feature attributes and large-scale data sets. In view of this, this paper proposes an analysis method that aims to define spatio-temporal graph association rules suitable for the Internet of Vehicles environment, and an abnormal behavior mining algorithm designed for spatio-temporal characteristic attributes. Through this method, not only can deep-level spatio-temporal rules be discovered from multi-dimensional data, but also dy- namic changes in the communication mode of the Internet of Vehicles can be dealt with. The main contributions are summarized as follows:

1) This paper integrates dynamic spatio-temporal graphs with the spatio-temporal characteristics of the IoV environment to precisely track the communication behav- iors of vehicle nodes at different times and locations, along with their changes. The core of this method is that it can comprehensively capture key information such as the vehicle's location, time, communication mode, and communication volume, thereby providing a com- prehensive analysis perspective for abnormal behavior node mining.

2) This paper proposes a spatio-temporal graph association rule mining algorithm. By defining spatio- temporal graph association rules, this algorithm is capable of mining deep associations within multidi- mensional data, particularly considering the spatio- temporal attribute characteristics of vehicle communi- cation in the IoV environment. It further delves into specific spatio-temporal features of vehicle communi- cation, such as communication frequency and commu- nication volume, providing a richer and more detailed informational dimension for anomaly detection.

3) Following an evaluation of the performance and efficiency of the dynamic spatio-temporal graph analy- sis method and spatio-temporal graph association rule mining algorithm proposed in this study, two datasets (NSL-KDD and SPMDD) were chosen as the basis for experiments. This experimental design aims to thor- oughly verify the effectiveness and mining efficiency of our previously proposed algorithm in practical appli- cations. The core focus of the experiments is on com- paring the execution times of our research algorithm with other centralized data mining algorithms, while also observing their performance in generating frequent rule numbers and how they are affected by different minimum support thresholds.

The structural arrangement of this article is as follows: In Section 2, we comprehensively review the cutting-edge research and related academic literature in the field of association rule mining, aiming to lay a solid theoretical foundation and methodological framework for this research. The following section three introduces in detail the spatio-temporal graph association rule mining method (ST-GARs) developed in this study, including the overall architecture of the algorithm and the key technical points to implement the method. Section 4 demonstrates the excellent performance of our proposed algorithm through experiments performed on two different data sets, thereby proving the effectiveness and application value of the algorithm. Finally, Section 5 concludes this study and the future work.

2. Related Work

2.1. Association Rule

Agrawal et al. [12]first proposed the concept of association rule mining, pioneering the application of association rules in commercial databases. Association rule mining [13–15] is a key technique in data mining, aimed at discovering interesting relationships between variables in large datasets. It helps researchers understand the potential associations in data by identifying frequently occurring patterns, asso- ciations, or causal structures. With the growing need for time series data analysis, temporal association rule mining emerged [16]. Compared with traditional association rules, temporal association rules consider the time factor in the mining process and can identify data patterns that change over time[17]. This is especially important for systems where the state of data changes rapidly over time. For example, Ozdenet al. [18] proposed sequence patterns and association rules considering the time dimension, adding dynamic con- text to the data for more refined analysis. Nasr M[19]et al. introduced Classification Association Rules (CARs), further segmenting the domain of association rule mining. Liu et al. [20] proposed an algorithm for mining association rules in sparse time series data common in industrial systems. Zhan et al.[21] proposed a new definition of tense correlation. Wang et al. [22] proposed a novel framework for mining temporal association rules, discovering rule sets through frequent rule trees. Although the proposed methods are useful, they do not reflect spatial attribute information.

In the realm of mobile edge computing , DONG et al. [23] have introduced a novel task offloading method based on the framework of prospect theory. Their study leverages behavioral economic principles to optimize decision-making in task offloading, effectively enhancing computational efficiency and reducing energy consumption. This research demonstrates the application of behavioral economics in managing computational tasks, aiming for improved system performance. Zhang et al. [24] have developed a new method for vehicle cooperative communication, integrating fuzzy logic with signal game strategies. This approach utilizes fuzzy logic to handle uncertain and incomplete information, while signal game strategies aid in making optimal decisions within complex communication environments. The introduction of this method offers a fresh perspective and technological pathway for addressing communication issues in the IoV. Furthering innovations in IoV. Wang et al. [25] have explored a novel edge caching approach using multi- agent deep reinforcement learning. This method employs deep learning technologies to autonomously learn optimal caching strategies, significantly increasing data processing speeds and reducing latency, thereby providing robust tech- nical support for data management in intelligent transporta- tion systems. Additionally. LI et al. [26] have proposed an energy-balanced routing method for wireless sensor net- works, centered on a forward-aware factor. This method considers the energy consumption during data transmission and the lifespan of the network, optimizing routing choices to balance energy usage and extend the network's overall longevity.

Temporal association rules can be used to analyze the patterns of change in vehicle behavior and communication over time, such as predicting traffic flow changes within specific time periods or identifying the types of traffic accidents most likely to occur during certain times. However, the IoV scenario poses additional challenges to temporal association rules. First of all, the Internet of Vehicles generates a huge amount of data with diverse data dimensions, including time, location, communication mode and communication volume, etc., which brings huge challenges to data processing and rule mining. Despite the great potential of temporal asso- ciation rules in the IoV, existing research still falls short in dealing with the high dimensionality and dynamics of IoV data. Current methods often rely on simplified data models or assumptions, which may lead to insufficient accuracy and efficiency when applied in actual IoV environments. Moreover, most existing algorithms fail to fully utilize the spatio-temporal characteristics of IoV data, limiting their application in dynamic environments.

2.2. The Application of Graph Models in Dynamic Networks

As the use of graph data becomes increasingly com- mon, researchers have developed a keen interest in mining static graph data. Many algorithms have been proposed for mining interesting graph patterns[27–29], such as frequent subgraphs. Moreover, a dynamic graph, or a sequence of graphs, incorporates time information. Patterns for mining tasks in dynamic graphs have been defined. The summa- rization and condensed representation of dynamic graphs have been studied. To reduce the complexity of dynamic graphs. A dynamic graph, or a graphsequence, takes the time information into account. Patterns are defined for dynamic graphs for mining tasks[30]. Dynamic graph summarization [31]and condensed representations of changes in dynamic graphshave [32] been investigated in order to reduce the complexity of the graph. Halder et al[33]. proposed a single pass supergraph based periodic pattern mining SPPMiner technique to identify periodic patterns in dynamic networks. These patterns are of great use. However, patterns in dy- namic graphs do not involve the attributes of vertices, hence they cannot describe temporal associations in multi-attribute graph sequences.

The temporal graph data structure[34] plays a central role in the analysis of IoV data due to its powerful capability to represent relationships. Graph pattern analysis can reveal complex interactions between nodes (such as vehicles, sensors). Techniques like frequent subgraph mining, extensively applied in the analysis of static networks, are now facing new challenges in dynamic networks, especially in highly dy- namic environments like the IoV.The dynamism of the IoV is reflected in the rapid changes of network topology and the real-time updating of node states. Traditional graph model mining methods cannot effectively deal with this dynamism. In recent years, researchers have proposed various graph model mining algorithms suitable for dynamic networks. For instance,Wu [35] introduced a frequent subgraph mining method for dynamic networks that considers changes in the graph model over time. Additionally, graph pattern analysis for spatio-temporal networks has also gained attention. For example, Wang and colleagues[36] introduction of Graph Pattern Association Rules (GPARs) has made substantial progress in the analysis of socialnetworks.

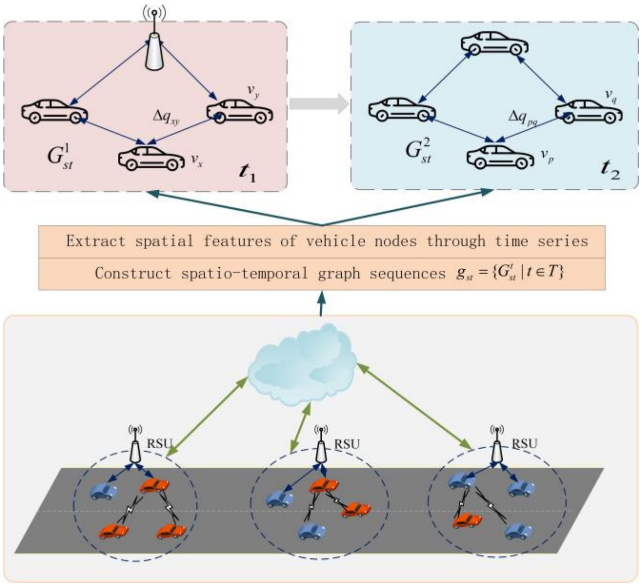


Figure 1. Construct spatio-temporal graph

3. Spatio-temporal Graph Association Rules Mining Algorithm (ST-GARs)

3.1. Spatio-temporal Graph Association Rule Mining

3.1.1. Spatio-temporal Graph Association Rule Mining

Table 1. Notations used in this paper

Notation	meaning
G_{st}	spatio-temporal graph
g_{st}	spatio-temporal graph sequence
V	vehicle node set
E	edge set
τ	time delay
Δt	time interval
R_{st}^q	spatio-temporal association rule set
Δq	message sending volume threshold
X	prerequisite event
Y	subsequent events
U	total number of events
$Sup(X \rightarrow Y)$	support
$Conf(X \rightarrow Y)$	confidence
q_{xy}	v_x to v_y information sent during a period
G_{st}^{α}	spatio-temporal graph in timeα
Z_{α}	Items of the space-time diagram G_{st}^{α}
Q	communication volume
F	Frequency of communication
f	communication volume threshold
Δq	communication frequency threshold

The spatio-temporal graph model is a complex network model that two dimensions of time and space to the traditional graph model, and used to represent and analyze entities and their relationships over time and space. This model is particularly suitable for describing entities whose locations and connection statuses change over time and space. In the Spatio-temporal graph model, nodes represent entities, edges represent relationships between entities, and each node and edge are associated with one or more timestamps, indicating the existence of an entity's state or relationship at specific times.

In the IoV environment, spatio-temporal graphs not only reflect the dynamic connectivity relationships between vehicles but also integrate the two key dimensions of time and space. Therefore, this paper proposes the following more accurate definition:

Definition 1 (Spatio-temporal Graph): A Spatio-temporal graph is a quadruple $G_{st}=(V,E,T,S)$, where $V=\{v_1,v_2,...v_i\}$ represents the set of nodes corresponding to vehicles or devices, each node representing a unique vehicle or device. $E=V\times V$ represents the set of directed edges between nodes, indicating the communication behavior between vehicles. T is the time dimension, corresponding to timestamps of activities for each edge or node. S is the space dimension, representing the geographic location of each node at specific times.

Definition 2 (Spatio-temporal Graph Sequence): A Spatio-temporal graph sequence is a series of Spatio-temporal graphs arranged in chronological order, where $G_{st}=(V,E,T,S)$ is the Spatio-temporal graph at time tt, and TT is the set of time points within the entire observation period. The Spatio-temporal graph sequence serializes the dynamic changes of Spatio-temporal graphs, providing a temporal perspective for analysis.

Definition 3 (Spatio-Temporal Graph Association Rule): A Spatio-temporal graph association rule indicates that under the condition of a given time delay δδ, if event AA occurs, then event BB is also likely to occur at time δδ.

Considering the Spatio-temporal graph association rules in the context of vehicle-to-vehicle communication in the Internet of Vehicles (IoV), especially focusing on the change in the volume of information sent over a certain period, we need to further refine and expand Definition 3. Such rules can not only capture the patterns of communication behavior between vehicles but also consider the quantitative characteristics of communication, such as increases or decreases in the volume of information sent.

Quantitative Spatio-temporal graph association rules defined in this way can not only capture the Spatio-temporal patterns of vehicle communication behavior but also quantify the trends in communication activities, thus providing a solid foundation for network anomaly mining tasks in the IoV.

As shown in Figure 1, the quantified spatio-temporal graph association rules defined in this way can not only capture the spatio-temporal patterns of vehicle communication behaviors, but also quantify the changing trends of communication behavior and build a Spatio-temporal graph model of the Internet of Vehicles, thus providing sufficient premise for the network anomaly mining task of the Internet of Vehicles.

3.1.2. Metrics for Spatio-temporal Association Rule Mining

(1) Spatio-temporal association rules have two attributes: frequency and credibility. The frequency of a rule indicates how frequently the rule appears, while the credibility indicates the probability of the conclusion appearing at the timestamp after the timestamp when the premise should appear. However, if we mine all Spatio-temporal association rules, we may get some infrequent or untrustworthy rules. Therefore, we need to distinguish between frequent and credible rules, which are considered important. To this end, two saliency measures for evaluating the importance of rules are presented. Support measures frequency, while confidence measures trustworthiness.

Definition 4 (Support and Confidence of Space-Time Association Rules): Suppose we have a time-space association rule, where: A represents the premise of the rule, which can be a condition in time or space; B represents the conclusion of the rule, which can be time or space Spatial conditions, Δt representing timestamps.

Define event A: vehicle v_x sends q amount of information to vehicle v_y at time t;

Define event B: a message $q+\Delta q$ sent by vehicle v_x to vehicle v_y at time $t+\Delta t$. Based on these definitions, support is defined as:

$$Sup(X\rightarrow Y)=Sup(X\bigcup Y)=P(X\bigcup Y)=\frac{A\bigcap B\,|\,\Delta q}{U} \quad (1)$$

Confidence ($Conf(v_x \xrightarrow{\Delta t} v_y)$): Confidence measures the credibility of the rule, that is, the probability that the conclusion will appear if the premise appears. Confidence can be expressed by the following formula:

$$Conf(X\rightarrow Y)=Sup(X\bigcup Y)/Sup(X)=P(Y\,|\,X)=\frac{A\bigcap B|\Delta q}{A} \quad (2)$$

Therefore, if we want to calculate the confidence and support of changes in the sending volume of two nodes in different time periods, we need to collect the frequency of these events and occur in the space-time graph sequence, and how often they occur together. Once you have this data, you can calculate it using the formula above. In practice, this often involves extracting this data from telematics communication logs and then applying the above formula to evaluate the strength of the rules.

(2) Support vector regression: The support vector machine is a pattern recognition method based on statistical theory. SVR is the application of support vector in the field of functional regression and can predict time series through regression analysis[38]. Consider a set of data point $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ in which $m(x_i, y_i \in R^n, i = 1, \dots m)$ is the total number of sample data. A linear regression function is formatted as:

$$f(x) = \omega^T \Phi(X_i) + b. \tag{3}$$

The weight vector ω and the bias value b are estimated by minimizing the following function:

$$min[\frac{1}{2}||\omega||^2 + C\sum_{i=1}^{m} \xi_i + \xi_i^*]. \tag{4}$$

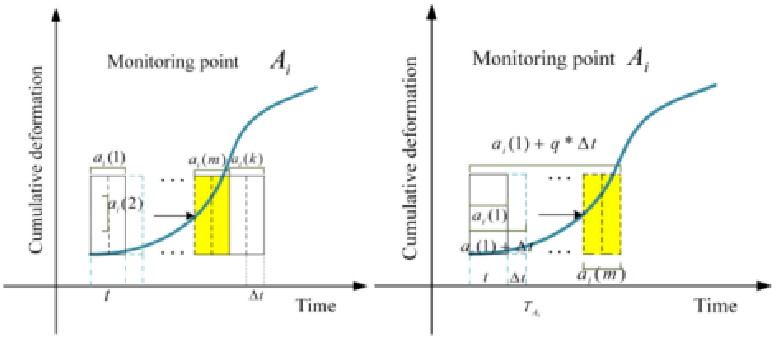
$$s.t. \begin{cases} (\omega x_i + b) - y_i \le \varepsilon + \xi_i \\ y_i - (\omega x_i + b) \le \varepsilon + \xi_i \\ \xi_i, \xi_i^* \ge 0, \varepsilon \ge 0 \end{cases} \tag{5}$$

The optimization problem can be expressed as the nonlinear regression function:

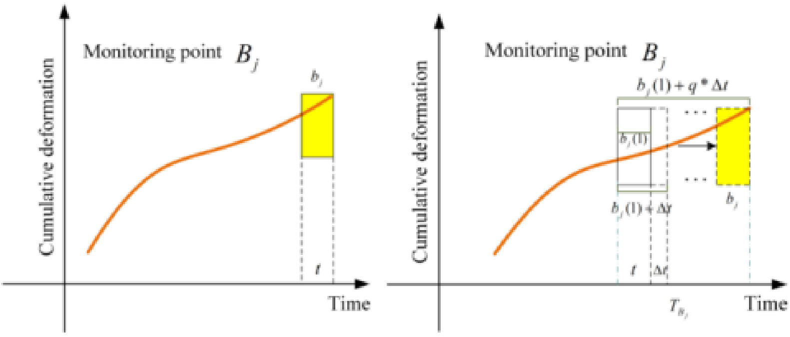
$$f(x) = \sum_{i=1}^{l} (\alpha_i - \beta_i) K(x_i, x) + b. \tag{6}$$

$$s.t. 0 \le \alpha_i \le C, 0 \le \beta_i \le C. \tag{7}$$

where α_i and β_i are Lagrange operators, $K(x_i, x) = (\Phi(x_i), \Phi(x))$ is the kernel function.



(a) Determination of relational time windows



(b) Mining the spatio-temporal association rules

Figure 2. Diagram of spatio-temporal association rule mining: (a) Determination of relational time windows; (b) Mining the spatio-temporal association rules

3.2. Network Communication Feature Extraction

Quantitative spatio-temporal association rules: $R_{st}^q : X \Rightarrow Y|_\tau^{\Delta q}$ indicate that under the conditions of a given time delay τ and information change threshold Δq, if the event X occurs, the event Y may also occur. Here X and Y are events defined on the attributes of nodes or edges respectively, and τ is the time interval, and Δq is the change threshold of the amount of information sent.

In the context of the IoV, event X can represent the communication activities of a group of vehicles at time t, and event Y represents the communication activities of this group of vehicles at time $t+\tau$. The information volume change threshold Δq is an indicator of the change in communication activity between X and Y, which can be an absolute quantity change or a relative proportion change. For example, if we define the event X as vehicle A sending a certain amount of data to vehicle B at time t, and the event Y is vehicle A sending more data to vehicle B at time, then it is the increment of the amount of data sent between the two communications. If this pattern occurs frequently during the observation period, then this quantitative Spatio-temporal association rule indicates that there is a trend in vehicle B's communication pattern to increase the amount of information sent over time. Specific to the analysis of communication traffic between vehicles, the indicators can be further accurately quantified Δq.

Information volume change threshold Δq : For vehicle communication between two time points t and $t+\tau$, $\Delta q = q_{t+\tau} - q_t$, where q_t and $q_{t+\tau}$ represent the amount of information sent at time t and $t+\tau$ respectively.

This article mainly considers abnormal points from two communication characteristics, namely communication frequency and communication volume. Communication frequency F and traffic volume Q are closely related to abnormal behavior. For example, if the communication frequency of a pair of vehicles suddenly increases in a short period of time, and the communication volume is abnormally large, this may imply an information flooding attack or data tampering behavior. This anomaly is detected by setting a threshold sum, that is, when $F > f$ or $Q > \Delta q$, we believe this may be unusual behavior. This paper determines these thresholds by statistically analyzing past communication data. Assuming that we have historical data of communication times between vehicles over a period of time, we can set thresholds f and Δq by calculating the statistical properties of these data.

The threshold of communication frequency can be determined by calculating the mean and standard deviation of historical communication frequency. The threshold can then be set to the mean plus times ω the standard deviation, where ω is a coefficient chosen to adjust the stringency of the threshold:

$$f = \mu_f + \omega \cdot \sigma_f$$

Here, the value of can be adjusted according to actual needs. In practice, it is often used to identify outliers that deviate significantly from the normal range.

Similarly, the traffic threshold Δq can be calculated by analyzing historical traffic data. First, calculate the mean and standard deviation of historical traffic. The threshold can then be set to the mean plus ω times the standard deviation:

$$\Delta q = \mu_q + \omega \cdot \sigma_q$$

Such calculations assume that abnormal communication behavior causes communication frequency or traffic volume to be significantly higher than normal. By setting the above thresholds, the system can automatically detect and alert potential abnormal behaviors that exceed the thresholds.

By simplifying the Internet of Vehicles communication logs in the data set into the instance set shown in Table 2, and calculating the thresholds of communication frequency and communication volume based on this set, as well as the support and confidence between nodes based on these thresholds.

Table 2. Communication log example

Event ID	Send ID	Received ID	Timestamp	Comm frequency	Comm volume
A	v_1	v_2	t_1	12	250kb
B	v_1	v_3	t_2	15	300kb
C	v_2	v_1	t_3	10	180kb
D	v_2	v_3	t_4	8	400kb
E	v_1	v_2	t_5	11	220kb
F	v_3	v_1	t_6	4	240kb

Table 2 showcases a simplified example of a communication log within a dataset. Based on the formulas provided, the average communication frequency is calculated to be μ_f =10; the standard deviation of communication frequency is μ_f=3.42. Assuming a threshold multiplier of ω=1, the communication frequency threshold is determined to be 13.42. Similarly, the total communication volume is calculated to be 265; the standard deviation of communication volume is 70; thus, the threshold for communication volume is set at 335. Following these thresholds, we can identify communication anomalies in events B (15 > 13.42) and D (400 > 335) as they exceed the set thresholds for both communication frequency and volume. For the calculations of support and confidence, given our defined specific spatio-temporal association rule for event D, where rule X represents vehicle A sending a communication to vehicle B at time T, and considering only 6 events in the example, with 2 exceeding the threshold (anomalies B and D), therefore $sup(R_{st}^t)$=$sup(v_1 \to v_2) = 2/6 = 1/3$, within the communications from v_1 to v_2, only one event meets the criteria, while there are a total of two communications (events B and D). Hence, the calculation is $conf(R_{st}^t)$=$conf(v_1 \to v_2)$=1/2 .

$$minsup = \mu_{sup(R_{st}^t)} + k \cdot \sigma_{sup(R_{st}^t)}$$

$$minconf = \mu_{conf(R_{st}^t)} + k \cdot \sigma_{conf(R_{st}^t)}$$

In conclusion, leveraging the computed thresholds for communication frequency and volume in the context of mining spatio-temporal association rules at nodes involves identifying frequent rules. This necessitates determining whether a rule is among those intended for mining. Therefore, a minimum support threshold $minsup$ and a minimum confidence threshold $minconf$ must be defined to categorize the computed rule set. A rule falls within the desired rule set, placed in the frequent rule set, when both the minimum support and minimum confidence criteria are met R . The calculation methods for minimum support and minimum confidence are illustrated as follows:

3.2.1. Mining Frequent Itemsets with Spatio-Temporal Graph Association Rules

Figure 4 shows the specific flow chart of the proposed algorithm. Its input is: a spatio-temporal graph $g_{st} = \{G_{st}^t \mid t \in T\}$, where $G_{st}^t = (V_t, E_t)$, the minimum support $minsup$,minimum confidence $minconf$, time constraints t, the communication volume Δq between nodes and the communication frequency f between nodes . The next step is to obtain frequent rules by generating candidate rule sets and then verifying their frequency by scanning $g_{st} = \{G_{st}^t \mid t \in T\}$. The common method for generating candidate rules is to produce frequent item vectors and then link each vector to form candidate rules. However, this method generates many candidate rules, resulting in low efficiency. The algorithm continuously and iteratively mines frequent rules, that is, it mines candidate (k-1)-rules based on the set of frequent k-rules, and verifies the frequency of candidate k-rules to obtain frequent k-rules. The iteration stops when no new frequent k-rules appear. To reduce the number of candidate k-rules and improve algorithm efficiency, this paper introduces a method of connection and correlation measurement as shown in the figure. Through the given inputs, it details the two steps of the proposed algorithm: discovering frequent rules and verifying the credibility of frequent rules.

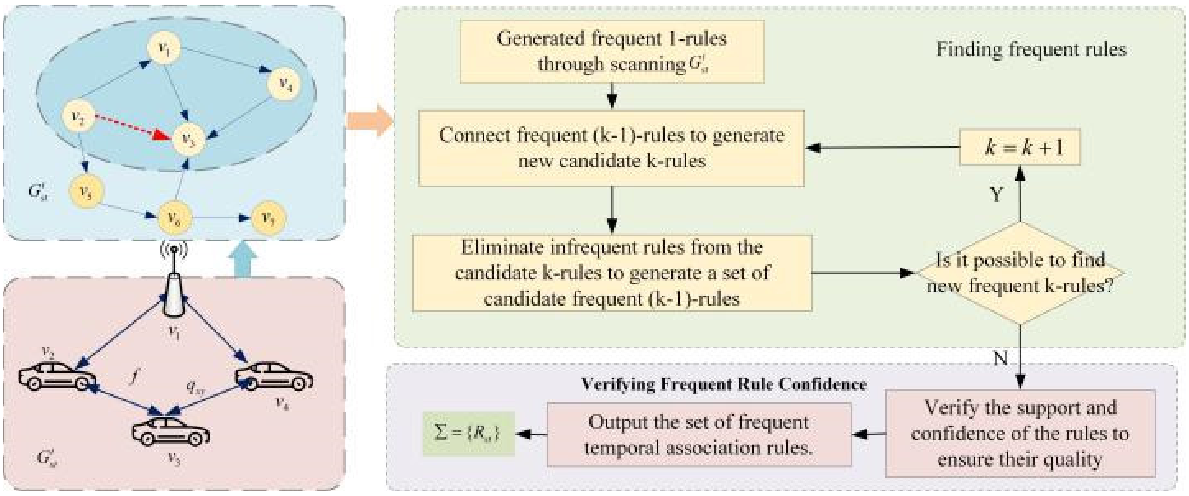


Figure 3. Algorithm flowchart

This stage involves generating candidate rules, followed by scanning the dataset D and calculating abnormal behavior events based on the communication frequency and volume among vehicles in the dataset. From these abnormal behavior events, frequent rules between vehicle nodes are calculated to identify outliers. The most significant time expenditure in this process occurs in scanning D to calculate support, hence the goal is to minimize the number of generated candidate rules. The conventional method of generating candidate rules includes first generating frequent itemset vectors, then linking every two frequent itemset vectors to form a candidate rule, but this method generates a large number of candidate rules, leading to low efficiency. The proposed algorithm mines frequent rules in an incremental and iterative manner, i.e., it mines candidate k-rules based on the existing frequent (k-1)-rule set and identifies frequent k-rules by verifying the frequency of these candidate k-rules. The iterative process ends when no new frequent k-rules are discovered. To reduce the number of candidate k-rules and improve algorithm efficiency, we specifically designed two new strategies: connection and pruning.

Firstly, we construct an item Z for each edge in $g_{st} = \{G^t_{st} \mid t \in T\}$. Given an edge $e_{ij} = (v_i, v_j)$, its constructed item behavior is $Z_{\alpha} = [((v_x, v_y), f, \Delta q) | e_{xy} \in E_{\alpha}]$, where Δq is the communication volume threshold within this spatio-temporal graph sequence, and f is the frequent communication threshold of the spatio-temporal graph sequence. From each item Z_{α} two rule relationships, Z→ $\varnothing$ and $\varnothing$ →Z, where $\varnothing$ represents a null graph, these rules signify the association of the existence or non-existence of an edge for a single node. Secondly, by scanning the spatio-temporal graph $g_{st} = \{G^t_{st} \mid t \in T\}$, we calculate the support of each 1-rule. Support refers to the frequency of occurrence of the edge corresponding to the rule node. We compare the calculated support with the minimum support threshold, filtering out frequent 1-rules that meet the criteria. Next, we generate candidate k-rules from any frequent (k-1)-rules. We connect two frequent (k-1)-rules using a connection method, producing a potential frequent k-rule. The iteration stops when no frequent k-rules appear. Ultimately, all frequent rules are identified.

The specific steps for implementing the algorithm can be divided into two parts: connection and pruning:

(1) Connection: In data mining algorithms, especially for mining rules from spatio-temporal graphs, the connection step plays a crucial role.

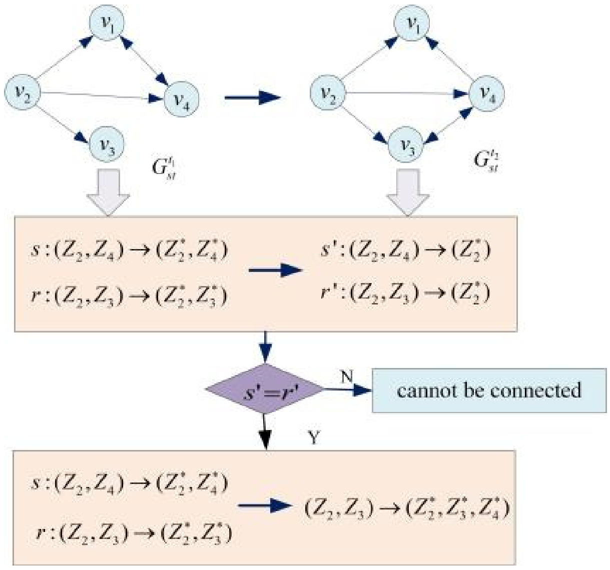


Figure 4. Connection process

The algorithm proposed in this paper utilizes this step to construct potential frequent k-rules. This process is based on the combination of two frequent (k-1)-rules, aiming to significantly reduce the search space and improve the efficiency of rule mining. Before detailing the connection process, it is essential to define the structure of a rule. A k-rule can be seen as a sequence composed of k items, where each item represents an attribute of an edge or vertex in the graph. In connecting two frequent (k-1)-rules to generate a new k-rule, we focus on the last item of the rules, which acts as a "bridge" in the connection process.

Specifically, the connection operation considers two frequent (k-1)-rules, and combines them into a new k item candidate rule only if these two rules have the same preceding (k-2) items, and their last items are different. This method ensures that the newly generated rules are logically coherent and avoids the generation of duplicate and irrelevant rules, thereby improving the overall efficiency of the algorithm. Furthermore, considering the dynamic nature of nodes in the spatio-temporal graph, i.e., vertices may transmit different amounts of information at different times Δq, this algorithm specially considers this dynamic feature.

The connection process can be described as follows: Given two frequent (k-1)-rules, r and s, if the first (k-2) items of r are the same as the first (k-2) items of s, but their last items are different, then the first (k-2) items of r can be combined with the last item of r and s, forming a new k-rule. The validity of this new rule will be determined by subsequent support and confidence tests. The specific steps are as follows:

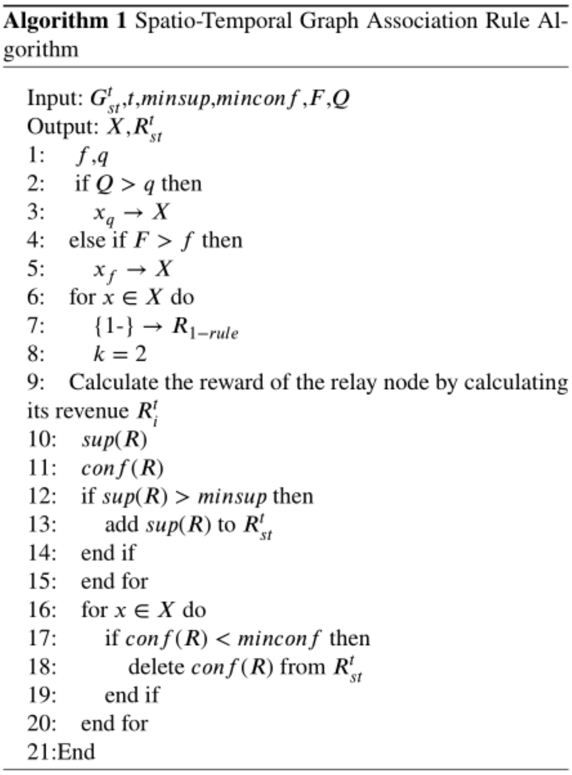
Step 1: Evaluate whether two sub-rules are consistent. For any rule in the frequent (k-1)-rules, s and r, if the sub-rule obtained by removing the last item of s is the same as the sub-rule obtained by removing the last item of r, proceed to step two.

Step 2: Rule connection. If the subsequent part of r is not empty, and the last item of r meets the compatibility conditions with the subsequent part of s, then the last item of s is added to the subsequent part of r to form a new rule. Conversely, if the subsequent part of s is empty, but the last item of s meets the compatibility conditions with the premise part of r, then the last item of s is added to the premise part of r to produce a new rule.

(2) Pruning: As another key procedure in Step 1 of the proposed algorithm, the pruning procedure deletes infrequent k-rules from the k-rules gained after joining and generates the candidate k-rules. The pruning procedure deletes infrequent rules based on the idea that, according to Corollary 1, if any sub-rule of a k-rule s is infrequent, then s is infrequent. The procedure is that for each k-rule s gained after joining, if any (k-1)-rule (which is a (k-1)-rule) of s is not in r, then delete s. The remained k-rules are the candidate k-rules.

3.2.2. Verifying Frequent Rule Confidence

In the final phase of the analysis process, our goal is to ensure that the mined frequent rules are not only frequent but also highly credible. To achieve this, we need to verify the confidence of each obtained frequent rule to assess their practical application value. In this process, a rule $minconf$ is considered important and meaningful, and non-trivial only if its confidence level is not less than the predetermined minimum confidence threshold $minconf$. The confidence of the obtained frequent rules is verified to be not less than $minconf$. Only the non-trivial rules with a confidence level not less than $minconf$ are considered important non-trivial rules. This step not only enhances the practicality and interpretability of the rules but also helps avoid generating too many irrelevant or misleading rules, thereby improving the overall efficiency and quality of the analysis.



Algorithm 1 Spatio-Temporal Graph Association Rule Algorithm

```
Input: $G_{st}^{t}$,$t$,$minsup$,$minconf$,$F$,$Q$
Output: $X$,$R_{st}^{t}$
$f$,$q$
if $Q > q$ then
    $x_q \rightarrow X$
else if $F > f$ then
    $x_f \rightarrow X$
for $x \in X$ do
    $\{1\text{-}\} \rightarrow R_{1-rule}$
    $k = 2$
Calculate the reward of the relay node by calculating its revenue $R_i^t$
  $sup(R)$
  $conf(R)$
  if $sup(R) > minsup$ then
      add $sup(R)$ to $R_{st}^{t}$
  end if
  end for
  for $x \in X$ do
      if $conf(R) < minconf$ then
          delete $conf(R)$ from $R_{st}^{t}$
      end if
  end for
End
```

description for model deployment data collection. The selection of this dataset is due to its authenticity and diversity, which assists in validating the robustness and applicability of spatio-temporal graph association rule algorithms in real deployment scenarios.

4. Experimental Results

In the relevant chapters of this study, we designed a series of experiments to evaluate the performance of the proposed algorithm spatio-temporal graph association rule mining (ST-GARs). Specifically, in Section 4.1, we introduced the datasets used in the experiments and the configuration of the testing environment. Subsequently, in Section 4.2, we conducted a comparative analysis of the performance of the ST-GARs method with several classical data mining methods including traditional association rule mining methods [39] (ARs), frequent itemset mining methods [36](T-FS-tree), and the SOC algorithm on two real datasets. The focus of this comparison includes the algorithm’s runtime, the number of highly associated nodes it can mine, and the number of generated association rules. Through these metrics, we aim to comprehensively evaluate the efficiency and effectiveness of the ST-GARs method in practical applications, especially in handling large-scale IoV data, compared to traditional methods’ advantages and potential.

4.1. Datasets and Environment

NSL-KDD dataset: NSL-KDD [40]is a revised version of the renowned KDD’99 dataset, composed of four sub- sets: KDD Test+, KDD Test-21, KDD Train+, and KDD Train+_20%. Among them, KDD Test-21 and KDD Train+_20% are subsets of KDD Train+ and KDD Test+. Each record includes 43 features, with 41 features representing the traffic input itself, while the last two features denote the label (normal or attack) and score (severity of the traffic input).

The dataset encompasses four different types of attacks: Denial of Service (DoS), Probing, User to Root (U2R), and

Remote to Local (R2L). This provides a diversified experimental scenario that aids in assessing the performance of algorithms under various attack types.

Safety Pilot Model Deployment Data (SPMDD) dataset: This dataset was collected during the Safety Experiment Model Deployment, containing Basic Safety Messages (BSM), vehicle trajectories, and various vehicle information interaction data. These data primarily originate from onboard and roadside devices, offering a detailed environmental Subsequently, we analyzed the performance of these algorithms in terms of runtime and the impact of parameters. All experiments were conducted in Python 3.9 on a personal computer equipped with a 2.50 GHz Intel Core CPU and 4.0GB RAM. Each algorithm was run 20 times to calculate the average results.

4.2. Performance Evaluation

we used a three-dimensional scatter plot, as shown in Figure 5, to visually demonstrate the relationship between the confidence, support and correlation between nodes of association rules. It aims to conduct an intuitive visual analysis of data mining results in the Internet of Vehicles.

From the observation in Figure 5, it can be seen that most of the rules showing high confidence and support also show significant node correlation. This phenomenon reveals a key trend in the Internet of Vehicles communication data: frequent and highly credible. Interaction patterns of degrees are often closely related to close connections between nodes. Such phenomena further confirm the effectiveness of the spatio-temporal graph association rule method in identifying abnormal behaviors. Specifically, in the analyzed data sets, those rules with higher support and confidence often do not conform to normal traffic flow patterns or regular communication behaviors among vehicles. Nodes corresponding to these rules often show significant inter-correlation in the spatial dimension. Therefore, this three-dimensional scatter plot provides us with a powerful analysis tool, which can not only intuitively show the trend of correlation degree corresponding to the rules in the Internet of Vehicles data.

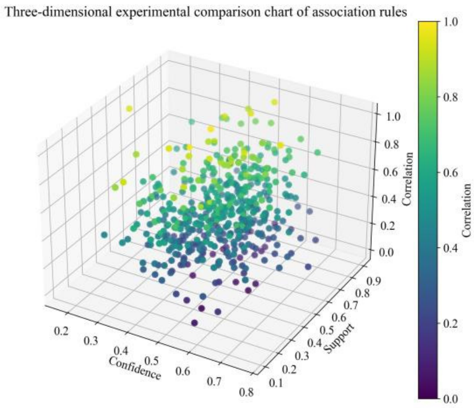


Figure 5. Visual analysis chart of correlation between Internet of Vehicles communication behavior

Experiment 1: In this experiment, we adopted a method of incremental dataset updates (updating 20% of the data each time) to compare the execution times of four methodologies: ST-GARs, ARs, T-FS-tree, and SOC. As demonstrated by the experimental results shown in Figure 6, the ST-GARs method exhibited the shortest runtime across two datasets as the communication volume of data increased. Notably, upon the dataset being updated to 100%, the execution time of ST-GARs reduced by nearly 50% compared to the traditional ARs algorithm. Moreover, as observed in Figure 6, the interval of required runtime displayed a trend of gradual reduction with the increase in data communication volume, indicating a progressive improvement in mining efficiency of ST-GARs over time. These findings distinctly highlight the significant time efficiency advantage of ST-GARs in processing large-scale vehicular network data, effectively enhancing the overall efficiency of mining.

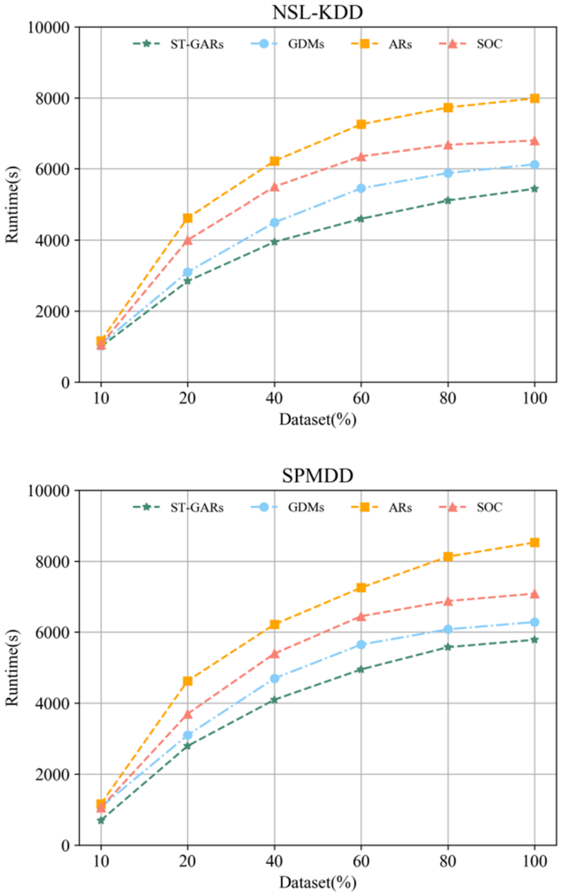
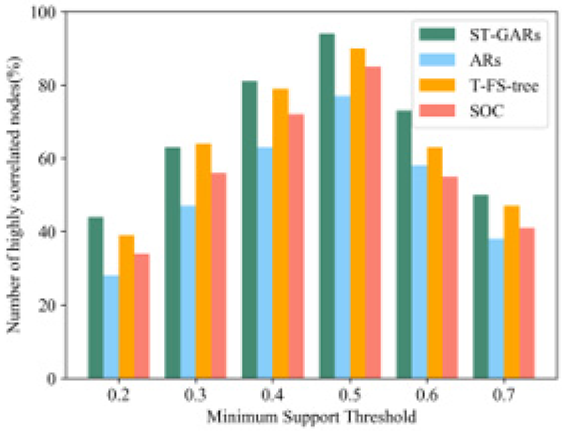


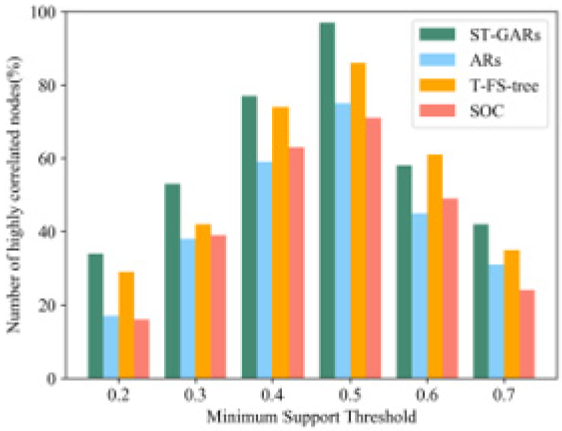
Figure 6. Compare the running times of different algorithms as the dataset is updated

Experiment 2: For this experiment, we compared the relationship between the number of highly relevant nodes and rules generated by different algorithms and the set minimum support threshold. The experimental results are shown in Figures 6 and 7 respectively. These results clearly show that compared with other algorithms, the ST-GARs algorithm proposed in this study can derive a larger number of high-correlation node set rules, especially when minsup=0.5, reaching an optimal state. The number of highly relevant node sets and rules mined by ST-GARs still exceeds that of other algorithms. This finding shows that the ST-GARs algorithm has significant advantages in min- ing capabilities, especially in revealing potential abnormal behavior information in the dataset.

Experiment 3: For this experiment, with the data size fixed, the execution time of different algorithms was compared by changing the value of from 0.2 to 0.7, where minsup=0.5. The results are shown in Figure 8 and 9. As we can see, in terms of time performance indicators, as the minimum support threshold increases, the running time of each algorithm first increases and then decreases. When minsup=0.5, several algorithms achieve the best state. Under the same minimum support threshold, ST-GARs performs better than other algorithms on both data sets. The closer the minimum support threshold is to 0.5, the higher the mining efficiency of ST-GAR is. In addition, this result also hints at the advantages of the ST-GARs algorithm in processing large-scale data. As the data set grows, algorithms that can effectively reduce the number of database scans will be more likely to maintain high performance. Therefore, the ST-GARs algorithm not only performs well on small-scale data sets, but is also very suitable for processing large and complex Internet of Vehicles data. In the future, we plan to further optimize the algorithm to adapt to a more dynamic and changeable Internet of Vehicles environment, while also considering the application potential of the algorithm in other fields, such as intelligent transportation systems and smart city management. Through continuous optimization and expansion, we believe that the ST-GARs algorithm can play a greater role in the field of spatio-temporal graph association rule mining.

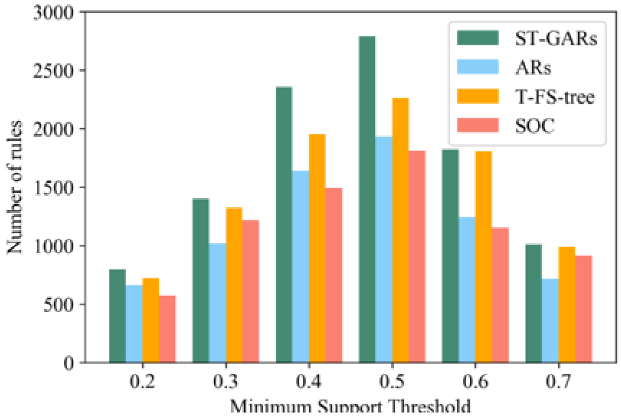


(a) NSL-KDD dataset

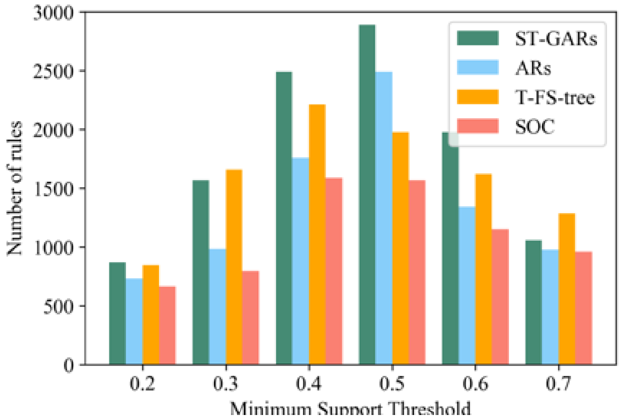


(b) SPMDD dataset

Figure 7. Comparison of the number of highly associated nodes mined from different datasets



(a) NSL-KDD dataset



(b) SPMDD dataset

Figure 8. Comparison of the quantity of rules extracted from various datasets

4.3. Discussion

Yu and colleagues [41] introduced a new Quantum- Genetic based OLSR protocol for Mobile Ad hoc Networks. This new protocol combines the advantages of quantum computing with the searching capabilities of genetic algo- rithms, aiming to enhance the efficiency of routing and net- work performance. GE et al. [16] proposed a new multi-hop clustering algorithm for Vehicular Ad Hoc Networks. This algorithm enhances data transmission and processing capabilities in vehicle networks by optimizing network structure and communication efficiency, thereby better addressing the communication demands in highly dynamic traffic environments. Zhang and others explored [42] a new UAV-assisted task offloading system, which employs the Dung Beetle Optimization Algorithm and deep reinforcement learning techniques. This system is designed to improve task offloading efficiency in edge computing environments, utilizing UAVs to optimize data processing and resource allocation decisions. Zhang and his team [43] introduced a new method for estimating missing data, using FNN- based tensor heterogeneous ensemble learning. This method specifically addresses data missing issues in the Internet of Vehicles (IoV), improving data integrity and enhancing the accuracy of data processing and analysis through advanced machine learning technologies.

In the field of communication and data management in the IoV, traditional spatio-temporal graph association rule mining is used to identify anomalous nodes, primarily rely- ing on historical data and the correlation of time and space to detect abnormal patterns. Firstly, the Quantum-Genetic based OLSR protocol, which combines quantum computing with genetic algorithms, optimizes routing choices, enhancing the network's self-organizing capabilities and efficiency, particularly effective in dynamically changing network environments. The multi-hop clustering algorithm improves the stability and efficiency of data transmission by optimizing communication links between vehicles, which is crucial for identifying anomalous nodes in the high-speed IoV environment. The UAV-assisted task offloading system, utilizing the Dung Beetle Optimization Algorithm and deep reinforcement learning, achieves resource optimization in complex edge computing scenarios, enhancing the flexibility and response speed of task processing. Lastly, the FNN-based tensor heterogeneous ensemble learning method focuses on addressing data missing issues, improving data integrity, thus making anomaly detection more accurate and reliable.Compared to spatio-temporal graph association rules, these methods demonstrate higher efficiency and accuracy in handling the dynamism and complexity of IoV data, better adapting to the rapid development and security needs of the IoV environment.

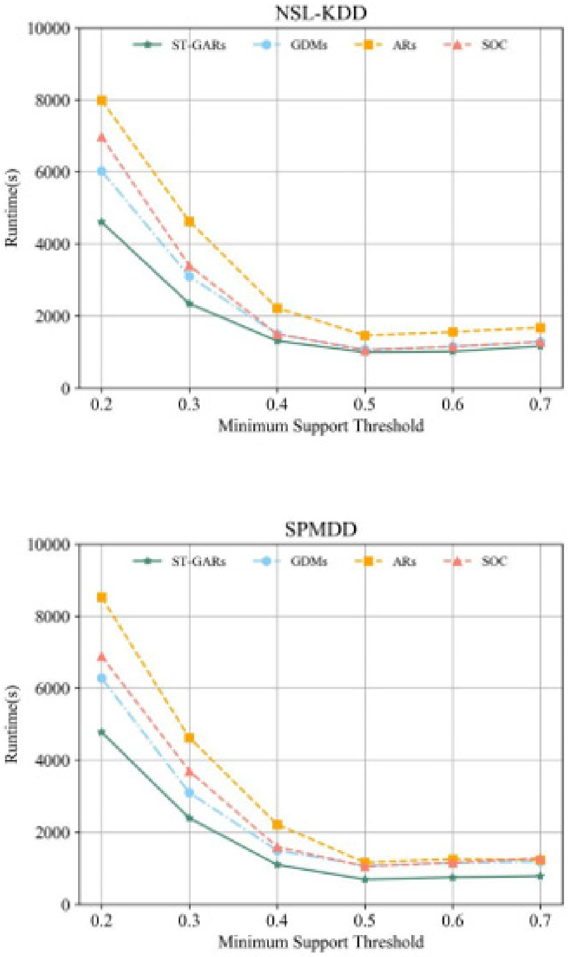


Figure 9. Comparison of the execution times of different algorithms on datasets

5. Conclusion and Future Work

This study introduces a novel spatio-temporal graph association rule mining algorithm tailored for analyzing vehicle communication behaviors in dynamic vehicle network environments. The algorithm excels at identifying nodes that frequently send or receive large volumes of data, aiding in the detection of information flooding, data leaks, or denial of service attacks. It operates in two main phases: initially identifying frequent temporal association rules, and subsequently verifying their credibility through innovative connectivity and correlation metrics. This approach has proven more efficient and accurate than existing methods in mining significant temporal association rules within vehicle networks.

Experimental tests on the NSL-KDD and SPMDD data sets demonstrate the superior effectiveness and efficiency of our algorithm in uncovering hidden temporal correlations in network communication behaviors among vehicles in the IoV. The results indicate a notable improvement in network security and the stability of the IoV system.

Future enhancements will focus on scaling the algorithm for larger and more complex IoV settings, extending its application to other networked environments like smart cities and industrial IoT, and improving its real-time capabilities to better adapt to dynamic conditions. These advancements are aimed at bolstering the overall security and reliability of the IoV through deeper analysis of communication dynamics and potential cybersecurity threats.

References

[1] Baofeng Ji, Xueru Zhang, Shahid Mumtaz, Congzheng Han, Chunguo Li, Hong Wen, and Dan Wang. Survey on the internet of vehicles: Network architectures and applications. IEEE Communications Standards Magazine, 4(1):34–41, 2020.

[2] Suvendu Kumar Nayak and Ananta Charan Ojha. Data leakage detection and prevention: Review and research directions. Machine Learning and Information Processing: Proceedings of ICMLIP 2019, pages 203–212, 2020.

[3] Chun-yan Zhang, Lin Zhao, Haotian Zhang, Meng-na Chen, Ru-yao Fang, Ying Yao, Qi-peng Zhang, and Qian Wang. Spatial-temporal characteristics of carbon emissions from land use change in yellow river delta region, china. Ecological Indicators, 136:108623, 2022.

[4] Fei Xiang, Xiaorong Wang, Xinliang He, Zhenghong Peng, Bohan Yang, Jianchu Zhang, Qiong Zhou, Hong Ye, Yanling Ma, Hui Li, et al. Antibody detection and dynamic characteristics in patients with coronavirus disease 2019. Clinical Infectious Diseases, 71(8):1930– 1934, 2020.

[5] Gowtham Atluri, Anuj Karpatne, and Vipin Kumar. Spatio-temporal data mining: A survey of problems and methods. ACM Computing Surveys (CSUR), 51(4):1–41, 2018.

[6] Qiwei Ma, Wei Sun, Junbo Gao, Pengwei Ma, and Mengjie Shi. Spatio-temporal adaptive graph convolutional networks for traffic flow forecasting. IET Intelligent Transport Systems, 17(4): 691–703, 2023.

[7] Zheyi Pan, Songyu Ke, Xiaodu Yang, Yuxuan Liang, Yong Yu, Junbo Zhang, and Yu Zheng. Autostg: Neural architecture search for predictions of spatio-temporal graph. In Proceedings of the Web Conference 2021, pages 1846–1855, 2021.

[8] Lubna Fayez Eliyan and Roberto Di Pietro. Dos and ddos attacks in software defined networks: A survey of existing solutions and research challenges. Future Generation Computer Systems, 122:149– 171, 2021.

[9] HN Lakshmi, Santosh Anand, and Somnath Sinha. Flooding attack in wireless sensor network-analysis and prevention. International Journal of Engineering and Advanced Technology, 8(5):1792–1796, 2019.

[10] Xiaoyan Liu, Feng Feng, Qian Wang, Ronald R Yager, Hamido Fujita, and José Carlos R Alcantud. Mining temporal association rules with temporal soft sets. Journal of Mathematics, 2021:1–17, 2021.

[11] Xubo Du and Fusheng Yu. A fast algorithm for mining temporal association rules in a multi-attributed graph sequence. Expert Systems with Applications, 192:116390, 2022.

[12] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining asso- ciation rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on Management of data, pages 207–216, 1993.

[13] Akbar Telikani, Amir H Gandomi, and Asadollah Shahbahrami. A survey of evolutionary computation for association rule mining. Information Sciences, 524:318–352, 2020.

[14] Federico Antonello, Piero Baraldi, Ahmed Shokry, Enrico Zio, Ugo Gentile, and Luigi Serio. A novel association rule mining method for the identification of rare functional dependencies in complex technical infrastructures from alarm data. Expert Systems with Applications, 170:114560, 2021.

[15] Akash Saxena and Vikram Rajpoot. A comparative analysis of asso- ciation rule mining algorithms. In IOP conference series: materials science and engineering, volume 1099, page 012032. IOP Publishing, 2021.

[16] Degan Zhang, Hui Ge, Ting Zhang, Yu-Ya Cui, Xiaohuan Liu, and Guoqiang Mao. New multi-hop clustering algorithm for vehicular ad hoc networks. IEEE Transactions on Intelligent Transportation Systems, 20(4):1517–1530, 2018.

[17] Yi Zhao. Discovery of temporal association rules in multivariate time series, 2017.

[18] Adnan Ozden, F Pelayo García de Arquer, Jianan Erick Huang, Joshua Wicks, Jared Sisler, Rui Kai Miao, Colin P O'Brien, Geonhui Lee, Xue Wang, Alexander H Ip, et al. Carbon-efficient carbon dioxide electrolysers. Nature Sustainability, 5(7):563–573, 2022.

[19] Mahmoud Nasr, Mohamed Hamdy, Doaa Hegazy, and Khaled Bah- nasy. An efficient algorithm for unique class association rule mining. Expert Systems with Applications, 164:113978, 2021.

[20] DongYue Liu, Bin Wu, Chao Gu, Yan Ma, and Bai Wang. A multidi- mensional time-series association rules algorithm based on spark. In 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pages 1946–1952. IEEE, 2017.

[21] Li Zhan, Fusheng Yu, and Huixin Zhang. A fast algorithm for mining temporal association rules based on a new definition. In 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pages 1548–1553. IEEE, 2017.

[22] Xin Wang, Yang Xu, and Huayi Zhan. Extending association rules with graph patterns. Expert Systems with Applications, 141:112897, 2020.

[23] De-Gan Zhang, Wen-Miao Dong, Ting Zhang, Jie Zhang, Ping Zhang, Gui-Xiang Sun, and Ya-Hui Cao. New computing tasks offloading method for mec based on prospect theory framework. IEEE Transac- tions on Computational Social Systems, 2022.

[24] De-Gan Zhang, Chen-Hao Ni, Jie Zhang, Ting Zhang, and Zhi-Hao Zhang. New method of vehicle cooperative communication based on fuzzy logic and signaling game strategy. Future Generation Computer Systems, 142:131–149, 2023.

[25] Degan Zhang, Wenjing Wang, Jie Zhang, Ting Zhang, Jinyu Du, and Chun Yang. Novel edge caching approach based on multi-agent deep reinforcement learning for internet of vehicles. IEEE Transactions on Intelligent Transportation Systems, 2023.

[26] Degan Zhang, Guang Li, Ke Zheng, Xuechao Ming, and Zhao-Hua Pan. An energy-balanced routing method based on forward-aware factor for wireless sensor networks. IEEE transactions on industrial informatics, 10(1):766–773, 2013.

[27] Vandana Bhatia and Rinkle Rani. Ap-fsm: A parallel algorithm for approximate frequent subgraph mining using pregel. Expert Systems with Applications, 106:217–232, 2018.

[28] Sidali Hocine Farhi and Dalila Boughaci. Two bi-objective hybrid approaches for the frequent subgraph mining problem. Applied Soft Computing, 72:291–297, 2018.

[29] Vijay Ingalalli, Dino Ienco, and Pascal Poncelet. Mining frequent subgraphs in multigraphs. Information Sciences, 451:50–66, 2018.

[30] Mostafa Haghir Chehreghani. Dynamical algorithms for data mining and machine learning over dynamic graphs. Wiley Interdisciplinary Reviews: Data Mining And Knowledge Discovery, 11(2): e1393, 2021.

[31] Sofia Fernandes, Hadi Fanaee-T, and João Gama. Dynamic graph summarization: a tensor decomposition approach. Data Mining and Knowledge Discovery, 32:1397–1420, 2018.

[32] Angelo Impedovo, Corrado Loglisci, Michelangelo Ceci, and Donato Malerba. Condensed representations of changes in dynamic graphs through emerging subgraph mining. Engineering Applications of Artificial Intelligence, 94:103830, 2020.

[33] Sajal Halder, Md Samiullah, and Young-Koo Lee. Supergraph based periodic pattern mining in dynamic social networks. Expert Systems with Applications, 72:430–442, 2017.

[34] Pietro Daverio, Hassan Nazeer Chaudhry, Alessandro Margara, and Matteo Rossi. Temporal pattern recognition in graph data structures. In 2021 IEEE International conference on big data (Big Data), pages 2753–2763. IEEE, 2021.

[35] Yan Li, Shuai Zhang, Lei Guo, Jing Liu, Youxi Wu, and Xindong Wu. Netnmsp: Nonoverlapping maximal sequential pattern mining. Applied Intelligence, pages 1–24, 2022.

[36] Xin Wang and Yang Xu. Mining graph pattern association rules. In International Conference on Database and Expert Systems Applica- tions, pages 223–235. Springer, 2018.

[37] Amol Kapoor, Xue Ben, Luyang Liu, Bryan Perozzi, Matt Barnes, Martin Blais, and Shawn O'Banion. Examining covid-19 forecast- ing using spatio-temporal graph neural networks. arXiv preprint arXiv:2007.03113, 2020.

[38] Xianglong Luo, Wenjuan Gan, Lixin Wang, Yonghong Chen, and Xue Meng. A prediction model of structural settlement based on emd-svr- wnn. Advances in Civil Engineering, 2020: 1–11, 2020.

[39] Bao Huynh, Lam BQ Nguyen, Duc HM Nguyen, Ngoc Thanh Nguyen, Hung-Son Nguyen, Tuyn Pham, Tri Pham, Loan TT Nguyen, Trinh DD Nguyen, and Bay Vo. Mining association rules from a single large graph. Cybernetics and Systems, 55(3):693–707, 2024.

[40] Tongtong Su, Huazhi Sun, Jinqi Zhu, Sheng Wang, and Yabo Li. Bat: Deep learning methods on network intrusion detection using nsl-kdd dataset. IEEE Access, 8:29575–29585, 2020.

[41] De-gan Zhang, Yu-ya Cui, and Ting Zhang. New quantum-genetic based olsr protocol (qg-olsr) for mobile ad hoc network. Applied Soft Computing, 80:285–296, 2019.

[42] Degan Zhang, Zhihao Zhang, Jie Zhang, Ting Zhang, Lei Zhang, and Hongtao Chen. Uav-assisted task offloading system using dung beetle optimization algorithm & deep reinforcement learning. Ad Hoc Networks, page 103434, 2024.

[43] Ting Zhang, De-gan Zhang, Hao-ran Yan, Jian-ning Qiu, and Jin- xin Gao. A new method of data missing estimation with fnn- based tensor heterogeneous ensemble learning for internet of vehicle. Neurocomputing, 420:98–110, 2021.