Journal of the Brazilian Computer Society, 2026, 32:1, doi: 10.5753/jbcs.2026.5664

Comprehensive Evaluation of Hybrid and XAI-Based Feature Selection for Intrusion Detection: A Smart City Perspective

Felipe N. Dresch [**Federal University of Pampa** | ***felipedresch.aluno@unipampa.edu.br***]
Felipe H. Scherer [**Federal University of Pampa** | ***felipescherer.aluno@unipampa.edu.br***]
Matheus M. Ciocca [**Federal University of Pampa** | ***matheusciocca.aluno@unipampa.edu.br***]
Vagner E. Quincozes [**Fluminense Federal University** | ***vequincozes@midiacom.uff.br***]
Silvio E. Quincozes [**Federal University of Pampa, Federal University of Uberlândia** | ***silvioquincozes@unipampa.edu.br***]
Diego Kreutz [**Federal University of Pampa** | ***diegokreutz@unipampa.edu.br***]

AI Horizon Labs, Federal University of Pampa.
810 Tiaraju Ave., Ibirapuitã, Alegrete, RS 97546-550, Brazil.

Received: 10 March 2025 • **Accepted:** 05 July 2025 • **Published:** 16 April 2026

Abstract. The expanding connectivity within smart cities has dramatically increased the attack surface, posing significant challenges for Intrusion Detection Systems (IDSs). A critical aspect of effective IDSs is the selection of relevant features to accurately identify potential attackers. Traditional feature selection methods, including filter-based (fast but potentially less accurate), wrapper-based (accurate but computationally intensive), and embedded (classifier-dependent) approaches, each present inherent limitations. Recent advancements propose alternative strategies, such as using Explainable Artificial Intelligence (XAI) algorithms to enhance filtering techniques, hybridizing filter and wrapper methods, and combining these strategies to optimize performance. However, a systematic evaluation of these novel feature selection methods within the context of diverse smart city environments remains largely unexplored. This work presents a comprehensive assessment of feature selection techniques for IDSs in multiple Smart City domains, including healthcare and transportation. Our analysis focuses on evaluating the trade-offs between classification performance and feature reduction achieved by hybrid (IWSHAP), metaheuristics (GRASPQ-FS) and XAI-based approaches (SHAP Ranking). The experimental results indicate that XAI-based methods achieve a favorable trade-off between dimensionality reduction and predictive performance, consistently preserving high F1-Scores (often exceeding 90%) while simultaneously reducing the feature set by substantial margins (e.g., over 90%). Although metaheuristic approaches can achieve superior feature reduction, they often require meticulous tuning to prevent performance degradation. This study underscores the potential of XAI-driven feature selection to enhance IDSs' effectiveness within complex Smart City ecosystems.

Keywords: Metaheuristics, Intrusion Detection System, Feature Selection, Explainable Artificial Intelligence (XAI)

1 Introduction

The evolution of smart cities, driven by the integration of information, communication, and control technologies within Cyber-Physical Systems (CPS), has enabled sophisticated automation and real-time decision making across critical sectors such as healthcare [Abbas *et al*., 2024], energy [Diaba *et al*., 2024], and transportation [Din *et al*., 2024; Dui *et al*., 2024]. By improving connectivity and fostering synergies among diverse systems, smart cities aim to improve urban efficiency, sustainability, and quality of life. However, this technological advancement significantly expands the attack surface, highlighting security as a paramount concern [Saber and Mazri, 2021; Demertzi *et al*., 2023].

Intrusion Detection Systems (IDSs) are fundamental to mitigating these threats, continuously monitoring CPS layers to detect malicious activities [Mitchell and Chen, 2014; Kayan *et al*., 2022; Stutz *et al*., 2024]. In the dynamic and complex environments of smart cities, network attacks can compromise critical processes, leading to severe disruptions in productivity and ecosystem security [Cherdantseva *et al*., 2016; Evancich and Li, 2016; Su, 2018; Demertzi *et al*., 2023; Awajan, 2023; Javeed *et al*., 2023; Kim *et al*., 2024].

Feature selection is a strategic step in optimizing IDSs performance [Bakro *et al*., 2024; Li *et al*., 2024; Turukmane and Devendiran, 2024]. By eliminating irrelevant features, this process accelerates data processing and enhances detection accuracy [Turukmane and Devendiran, 2024; Fang *et al*., 2024; Pesaramelli and Sujatha, 2024; Quincozes *et al*., 2020].

However, traditional feature selection methods, which include filter-based, wrapper-based, and embedded approaches, each present inherent limitations that can affect scalability, interpretability, or adaptability to complex data environments. Filter-based methods, while computationally efficient, often fail to account for feature interactions or their impact on the specific learning model, potentially leading to the selection of redundant features or missed optimal combinations. Wrapper-based methods, in contrast, are highly dependent on the chosen classifier and computationally intensive due to repeated model retraining, making them prone to overfitting and inefficient for large datasets. Lastly, embedded methods integrate feature selection directly into the model training,

resulting in classifier-dependent features that may not generalize well to other algorithms and offer less flexibility in identifying a minimal feature subset.

To address these shortcomings, alternative approaches have been proposed: (i) replacing filter methods with Explainable Artificial Intelligence (XAI) algorithms [Khani *et al*., 2024; Arreche *et al*., 2024]; (ii) hybridizing filter and wrapper methods [Moradkhani *et al*., 2015; Quincozes *et al*., 2024]; and (iii) combining these various strategies [Scherer *et al*., 2024a].

Despite recent progress in feature selection methodologies, a thorough evaluation of cutting-edge techniques tailored to the unique demands of smart city environments is notably absent. This research gap is critical, as smart cities require a nuanced equilibrium between minimizing feature sets and maximizing intrusion detection efficacy. Although hybrid and XAI-driven approaches demonstrate potential to improve selection efficiency, the absence of standardized benchmarks impedes the determination of optimal strategies for different operational contexts.

In this work, we present a comprehensive evaluation of feature selection methods for IDSs by systematically comparing multiple approaches across diverse smart city scenarios (*e.g.,* vehicular, healthcare), encompassing XAI-based, metaheuristic, and hybrid strategies. The evaluated methodologies include Incremental Wrapper Subset Selection (IWSS) [Bermejo *et al*., 2009], SHapley Additive exPlanations (SHAP) Ranking, IWSHAP, and Greedy Randomized Adaptive Search Procedure with Priority Queue for Feature Selection (GRASPQ-FS) [Quincozes *et al*., 2024], as well as their integrated variants. Our findings demonstrate that the efficacy of each approach is highly context-sensitive, necessitating tailored evaluations for specific application domains.

2 Background

In this section, we delve into the foundational concepts of smart cities and CPS, analyze their associated security threats and detection methodologies, and discuss the application of feature selection within Intrusion Detection Systems.

2.1 Smart Cities and Cyber-Physical Systems

Smart cities represent a paradigmatic evolution in urban management, leveraging cyber-physical technologies to promote greater operational efficiency, environmental sustainability, and quality of life. The systematic incorporation of connected devices and sensors allows for the continuous collection and analysis of data from different domains, enabling real-time decision-making and the optimization of essential services such as transportation, energy, health, and public safety. In this context, communication between distributed components plays a central role, enabled by a wide range of technologies and protocols. Among these, we have, as an example, the Controller Area Network (CAN), which is widely used in vehicle embedded systems and adapted in urban scenarios to enable efficient communication between sensors, actuators, and control units. Figure 1 illustrates a representative smart city scenario, highlighting the integrated interaction between multiple connected subsystems.

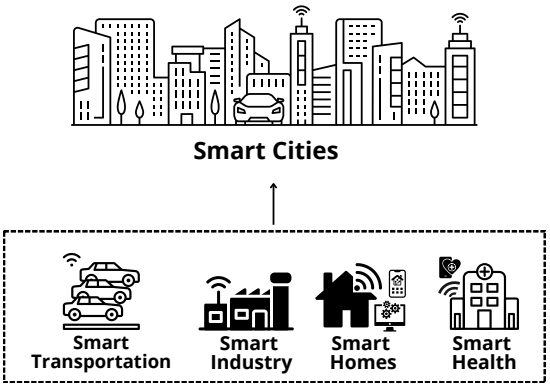


Figure 1. Smart City Applications.

CPSs amalgamate computational components, communication networks, and integrated hardware to interact with the physical environment in a coordinated and autonomous manner (*e.g.,* sensors and actuators). They embody the next generation of embedded systems, where the seamless integration and interaction between the physical and virtual realms are fundamental [Marwedel, 2021]. Unlike traditional systems, CPSs exhibit a profound integration between computational and physical processes, wherein computational components monitor and control physical processes, often through cyclic feedback, resulting in mutual influence.

The applications of CPSs within smart cities are diverse, spanning industrial environments where interconnected advanced manufacturing systems enhance production efficiency and flexibility; smart energy grids that integrate information and communication technologies into electrical networks to bolster energy supply efficiency and sustainability [Diaba *et al*., 2024]; transportation infrastructures, encompassing autonomous vehicles and real-time traffic management [Din *et al*., 2024; Dui *et al*., 2024]; medical and healthcare systems, such as the Internet of Medical Things (IoMT), which interconnect medical devices [Abbas *et al*., 2024]; and numerous other applications.

The archetypal architecture of a CPS comprises three primary layers: the perception layer, featuring sensors and actuators that interface directly with the physical environment; the transmission layer, responsible for the communication and transport of information between the perception and application layers; and the application layer, where data are processed for decision-making, control algorithm execution, and the provision of interfaces for users and other systems. The transmission layer is particularly critical, serving as the connective tissue that links the distributed components of the system, enabling real-time information exchange [Mahmoud *et al*., 2015; Kato *et al*., 2024; Faliagka *et al*., 2024].

Network connectivity is a quintessential characteristic of CPSs, facilitating the exchange of information among multiple distributed systems and devices. This connectivity is implemented through multiple communication technologies and protocols, encompassing both wired and wireless networks, as well as conventional and industrial standards such as the CAN network, and it supports a wide variety of interconnected devices.

2.2 Threats and Detection Methods

In smart city applications, security is paramount to ensure operational continuity, prioritizing service availability, data integrity, confidentiality, and authenticity [Ashibani and Mahmoud, 2017]. This emphasis stems from the critical nature of real-time monitoring and control inherent in the CPSs that constitute smart city infrastructure. To achieve these security objectives, we employ various cyber defense strategies, focusing particularly on the transmission layer. Although considered one of the most resilient layers, a successful breach permits attackers to manipulate, intercept, and retransmit information without constraints on the compromised channel, thereby endangering system operation and security [Cao *et al*., 2020]. Control of the transmission layer also facilitates network overload attacks, exposing CPSs to threats such as Denial of Service (DoS) and port scanning attacks [Cao *et al*., 2020; Sharafaldin *et al*., 2018].

To detect these malicious activities, we utilize IDSs that monitor and analyze network traffic for anomalous behavior. IDSs can leverage signature-based or anomaly-based techniques [Ahmad *et al*., 2021]. Signature-based techniques offer high accuracy for detecting known attacks with predefined signatures. Anomaly-based techniques, conversely, aim to identify anomalous behavior, including zero-day attacks, and generally offer faster detection speeds [Garcia-Teodoro *et al*., 2009; Kwon *et al*., 2019; Quincozes *et al*., 2020]. However, anomaly-based methods often produce a higher rate of false positives, as not all anomalies correlate with malicious actions.

Regardless of the technique employed, we recognize that an IDS must be configured with the most relevant features for the specific threat being targeted. Feature relevance directly correlates with the system's detection objectives. Therefore, we emphasize the careful selection of features during system preparation, as they dictate critical performance metrics such as identification time and accuracy.

To evaluate and refine identification and defense strategies, we understand the importance of characterizing attacks targeting diverse computer network and CPS contexts. In conventional networks, PortScan and Distributed Denial of Service (DDoS) attacks pose significant challenges. PortScan, a reconnaissance attack, involves identifying open ports to exploit as entry points for network compromise. DDoS attacks, on the other hand, aim to overload network resources, causing service interruptions and compromising critical system availability [Sharafaldin *et al*., 2018].

Within CPSs, each domain faces distinct attack vectors. For instance, fabrication and masquerade attacks are prevalent threats in vehicular CPSs. These attacks exploit vulnerabilities in CAN networks, which lack robust data integrity and authenticity mechanisms. Fabrication attacks involve injecting false information to manipulate processes, while masquerading allows attackers to impersonate legitimate users to gain unauthorized access. These attacks challenge identity and authentication controls, compromising data security and system trust [Jeong *et al*., 2024].

In the IoMT, security is paramount due to the direct impact on patient safety and privacy [Abbas *et al*., 2024]. Reconnaissance (Recon) and identity spoofing (Spoofing) are significant threats. Recon involves gathering detailed environmental information to map vulnerabilities, while spoofing falsifies device or user identities to manipulate data flow and gain unauthorized access. These attacks compromise data security, physical safety, and patient privacy, which are critical in sensitive medical environments [Dadkhah *et al*., 2024].

2.3 Feature Selection

The precise and appropriate selection of features processed by machine learning algorithms is crucial for IDSs to effectively identify diverse attack types, capturing their unique characteristics and patterns. An incorrect choice of features could not only diminish detection capabilities, but also significantly elevate the false positive rate, thereby compromising the system's efficiency and reliability in detecting specific threats. We recognize that this selection process can be executed through various methodologies, each tailored to specific domains.

2.3.1 Filter Methods

We understand that filter methods for feature selection utilize performance metrics independent of specific system algorithms, identifying optimal features based on a scoring algorithm followed by filtering. In this approach, we can rank features using metrics such as Information Gain (IG) and Information Gain Ratio (IGR). IG quantifies the reduction in uncertainty about a target variable given the value of an attribute, while IGR normalizes this measure to account for the attribute's cardinality, mitigating bias towards attributes with numerous categories [Esseghir, 2010]. Recent advances propose replacing traditional filter methods with XAI algorithms, as these algorithms also enable feature ranking based on their importance in predicting attack classes [Okada *et al*., 2025; Bataineh *et al*., 2024; E. L. Asry *et al*., 2024; Scherer *et al*., 2024a].

2.3.2 Wrapper Methods

We employ wrapper methods that integrate the machine learning algorithm into the feature selection process by evaluating feature subsets based on their predictive performance. These methods generate and assess multiple candidate subsets using the algorithm. We can utilize various strategies to construct these subsets. For instance, IWSS [Bermejo *et al*., 2009] constructs feature subsets by retaining performance-enhancing features and discarding noncontributing ones, yielding an optimized subset.

2.3.3 Embedded Methods

We acknowledge that the embedded methods perform feature selection concurrently with the execution of the detection algorithm, contrasting with other approaches where selection occurs offline. These methods aim to reduce the computational overhead associated with other selection techniques [Chandrashekar and Sahin, 2014; Guyon and Elisseeff, 2003].

2.3.4 Hybrid Methods

We utilize hybrid methods that combine the aforementioned approaches to identify an optimized feature subset with reduced computational effort. For example, Greedy Randomized Adaptive Search Procedure with Priority Queue for Feature Selection (GRASPQ-FS) employs a basic filtering algorithm to generate a concise set of candidate features, which are then used to build optimized solutions via a wrapper-based local search [Quincozes *et al*., 2020]. More recent hybrid methods incorporate XAI-based filtering to rank relevant attributes, followed by wrapper methods for optimal subset selection. For instance, IWSHAP [Scherer *et al*., 2024a] combines IWSS [Bermejo *et al*., 2009] with SHAP as an XAI-based filtering strategy.

3 Related Work

Smart cities present an increasing demand for increasingly accurate and optimized IDS models, capable of addressing both general scenarios and highly specialized domains (e.g., vehicle and healthcare environments). In this context, we have observed the development of various feature selection approaches, each employing a diverse range of methodologies and applications. The following discussion details these related works, which are summarized in Table 1 and compared with our approach.

We have observed that XAI-based tools have recently been introduced as filtering methods for feature selection, offering high precision in identifying the most relevant variables for specific models due to their explainable nature [Khani *et al*., 2024]. The IWSHAP method, for example, was proposed to optimize feature selection in an IDS tailored for CAN networks, maximizing learning model performance while ensuring decision interpretability, a crucial aspect in critical environments such as the vehicular domain. We recognize that the benefits of this approach are twofold: (i) it improves predictive performance metrics (*i.e.,* F1-Score, recall, precision, and accuracy), and (ii) it significantly reduces runtime and the number of features required to identify an attack [Scherer *et al*., 2024a].

We have also found that using XAI-based techniques for feature selection has gained prominence in other intrusion detection domains. The work [Arreche *et al*., 2024] proposes XAI-based approaches that utilize SHAP to assess the importance of features extracted from network traffic data. This research, conducted using the CICIDS-2017 [Sharafaldin *et al*., 2018] and RoEduNet-SIMARGL2021 [Mihailescu *et al*., 2021] datasets, demonstrates that XAI-guided feature selection can improve the effectiveness of the AI model in IDSs, as well as provide deeper insights into their decisions. Compared to traditional feature selection methods, we have found that XAI-based techniques showed superior performance, highlighting their potential in cybersecurity contexts. Moreover, the study underscores the importance of explainability in IDS analysis, allowing security analysts to interpret and act more precisely based on model outputs [Arreche *et al*., 2024].

Similarly, [Khani *et al*., 2024] applies XAI to improve the feature selection performance in IDSs. The study explores the applicability of SHAP to identify important features extracted from network traffic data. Using the CICIDS-2017 [Sharafaldin *et al*., 2018] and CIC-Darknet2020 [Habibi Lashkari *et al*., 2020] datasets, the research highlights that the selection of XAI-driven features not only improves the performance of the classification model, but also improves the understanding of the decision-making processes of these models. Comparative analyses with traditional feature selection methods reveal that XAI-based approaches achieve superior results, especially in classifying encrypted traffic.

Although [Khani *et al*., 2024] presents a promising application of XAI to enhance feature selection in IDSs, we believe that future research could explore additional avenues to further refine this approach. For example, incorporating metaheuristic algorithms could provide alternative strategies for navigating complex search spaces, while investigating hybrid feature selection methods that combine traditional techniques with XAI-driven approaches could complement the existing framework. These potential enhancements may offer further improvements in feature selection performance.

Similarly, [Sivamohan and Sridhar, 2023] introduces a novel approach using a Bidirectional Long-Short-Term Memory-based XAI framework (BiLSTM-XAI). This framework combines advanced neural network architectures with XAI techniques to improve both the detection accuracy and interpretability of intrusion behaviors. The paper presents a robust feature selection method through the Krill Herd Optimization (KHO) algorithm and reports a classification accuracy of 98.2% using the Honeypot and NSL-KDD [Tavallaee *et al*., 2009] datasets, showcasing its superior performance against existing detection methods. However, we have noticed that the dataset used in this study raises concerns due to its severe obsolescence, a critique that has been highlighted in the literature over the past decade [Divekar *et al*., 2018].

In [Shanbhag *et al*., 2024], the authors investigate various metaheuristic algorithms along with traditional machine learning classifiers to optimize feature selection in IDSs. Their approach incorporates techniques such as Genetic Algorithm, Particle Swarm Optimization, and Grey Wolf Optimization to enhance the performance of classifiers like Decision Trees and Random Forests. The effectiveness of these methods is assessed primarily through the test score metric, which accounts for multiple factors, including the F1-Score, recall, precision, recorded runtime, and the final subset length. Although we acknowledge that this study aligns with ours, demonstrating the potential of hybrid approaches to improve the detection accuracy of IDSs through metaheuristics, it does not explore the XAI-based approach. Furthermore, we have observed that the results obtained in the study rely on the obsolete NSL-KDD [Tavallaee *et al*., 2009] dataset.

Finally, we have seen that metaheuristics like GRASP have demonstrated significant efficiency in feature selection tasks, particularly in scenarios involving complex and heterogeneous datasets. GRASP, which combines a greedy approach with randomized search, has been adapted to optimize intrusion detection in the CPS perception layer by selecting features that improve classification accuracy while reducing computational overhead [Quincozes *et al*., 2020]. The integration of metaheuristics like GRASP with other filtering

Table 1. Related Works.
Cells Caption: ● Yes; ○ No.

Work	Selection Method	Scenario	XAI	Metaheuristics	Hybrid
Arreche *et al*. [2024]	Filtering	Educational Networks	●	○	○
Khani *et al*. [2024]	Filtering	Overlay Networks	●	○	○
Sivamohan and Sridhar [2023]	KHO	Traditional Networks	●	○	○
Shanbhag *et al*. [2024]	Metaheuristics	Traditional Networks	○	●	●
Quincozes *et al*. [2020]	GRASP-FS	Sensor Networks	○	●	●
Quincozes *et al*. [2024]	GRASPQ-FS	IEC-61850 Networks	○	●	●
Scherer *et al*. [2024a]	IWSHAP	CAN Networks	●	○	●
This Work	**IWSHAP, GRASPQ-FS, SHAP, IWSS**	**Smart Cities**	●	●	●

approaches has also shown promise. In [Quincozes *et al*., 2024], the combination of the GRASP-FS metaheuristic with a Priority Queue-based filtering approach proved to be highly effective in improving model precision and optimization.

4 Feature Selection Algorithms

In this section, we present a detailed analysis of the IWSHAP, SHAP Ranking, GRASPQ-FS, and IWSS algorithms, which constitute the methodological foundation of our study. We examine the underlying theoretical principles of each approach to provide a comprehensive understanding of their application in our research.

4.1 IWSS

We utilize the IWSS algorithm, which adheres to a wrapper methodology, employing an iterative and incremental approach for feature selection. To elucidate the IWSS process, we consider a dataset containing Y features as input. This process is illustrated in Figure 2 and is described below.

The process begins by sequentially iterating through each feature in the dataset, assessing the impact of its inclusion on the trained model's performance. At each iteration, the model performance metrics are computed. If the inclusion of a feature leads to an improvement in the evaluated indicators,

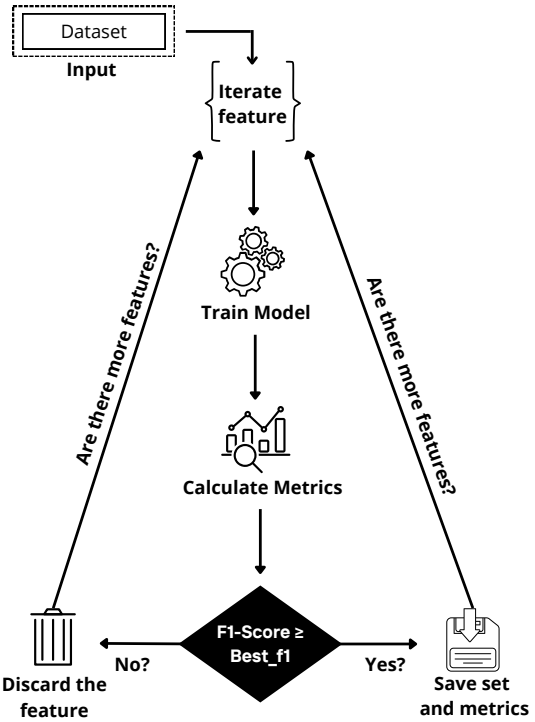


Figure 2. Illustration of the IWSS process.

we retain it in the selected subset. Conversely, if the feature does not contribute positively, we discard it and the process advances to the subsequent feature. This procedure continues

until all Y features have been analyzed.

Consequently, the algorithm yields a subset M of selected features, representing those that positively contributed to the model's performance [Bermejo *et al*., 2009].

4.2 IWSHAP

We employ IWSHAP, an extension of the IWSS wrapper-based algorithm. Specifically, IWSHAP is a hybrid method that leverages an XAI algorithm, namely SHAP, to rank features by their importance before applying the sequential feature analysis with IWSS. Once the features are ranked, we combine them incrementally, eliminating those that do not positively impact the F1-Score until the optimal set of features is identified, as illustrated in Figure 3.

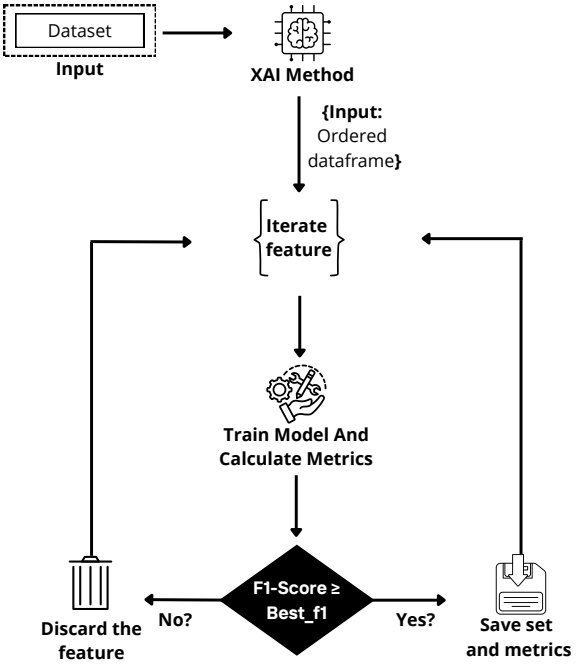


Figure 3. Illustration of the IWSHAP process.

This method is highly efficient, as it avoids blindly testing features and instead prioritizes the most relevant ones for detecting an attack. We have found that this approach has proven promising, reducing the number of features by more than 90% while achieving slight improvements in prediction performance metrics [Scherer *et al*., 2024a].

4.3 SHAP Ranking

We employ a ranking-based feature selection method that utilizes SHAP values to rank features by their importance. In practice, we can either select the top K features or apply a threshold cut-off based on the magnitude of the SHAP value. Both strategies aim to efficiently identify the most relevant features without an exhaustive search, although they carry the risk of misidentifying the optimal cutoff point.

To mitigate the risk of an arbitrary threshold, in this work, we perform an exhaustive search to determine the optimal cut-off that maximizes performance metrics. This systematic evaluation ensures a fair comparison between the SHAP Ranking method and other feature selection approaches. Figure 4 illustrates the complete process of the SHAP Ranking method, from feature classification to selection of the optimal set of features through an exhaustive search.

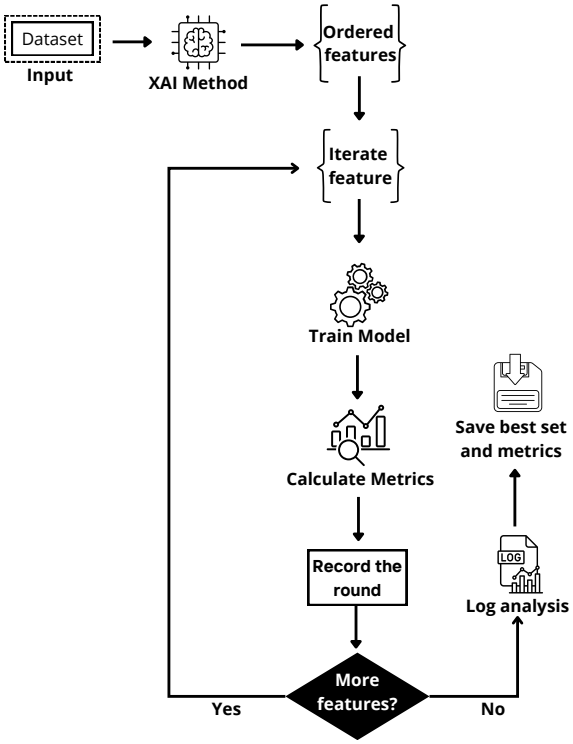


Figure 4. SHAP Ranking with exhaustive search for the best cutoff.

After ranking features using SHAP values, we implement an iterative cumulative procedure in which features are incrementally added to the model to evaluate their impact on performance. Initially, the model is trained using only the highest value feature; subsequently, it is re-trained with the highest value feature combined with the second highest-valued feature, and this process continues until all features are incorporated. Ultimately, we select the subset of k features that yields the highest performance metrics, such as F1-Score, recall, or precision [Scherer *et al*., 2024a].

However, it is important to note that this algorithm is significantly more resource-intensive than its alternatives, such as IWSHAP and IWSS, because it does not discard any features during the selection process.

4.4 GRASPQ-FS

We utilize GRASPQ-FS [Quincozes *et al*., 2024], which belongs to the class of hybrid approaches, similar to IWSHAP, and is an extension of the GRASP-FS metaheuristic [Quincozes *et al*., 2020]. The algorithm comprises two primary phases: the construction phase, where initial solutions are generated, and the local search phase, where these solutions are refined. The algorithm receives as input a labeled dataset, a user-defined number of features N to be selected, and stopping criteria for both phases. It also computes a Restricted Candidate List (RCL), built by ranking all features according to their Mutual Information (MI) and selecting the top-ranked ones. Figure 5 illustrates the step-by-step operation of the GRASPQ-FS algorithm, highlighting the construction and

local search phases.

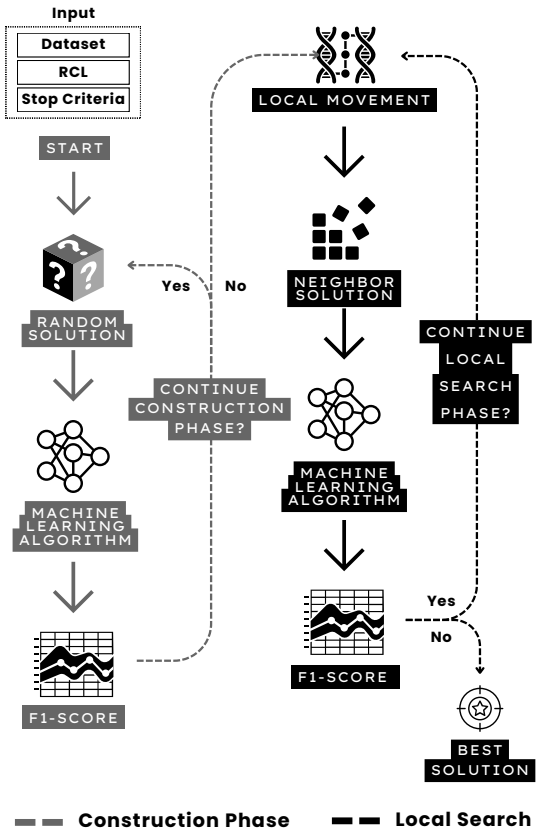


Figure 5. GRASPQ-FS process. Adapted from [Quincozes *et al*., 2024].

In the construction phase, GRASPQ-FS generates initial solutions by randomly selecting N features from the RCL. Each candidate solution is evaluated using a machine learning classifier, and its quality is measured via the F1-Score. The best-performing solutions are stored in a priority queue, which dynamically retains only the top candidates based on their F1-Scores. This selective mechanism limits the number of solutions passed to the next phase, improving computational efficiency while preserving detection performance. The left portion of Figure 5 represents this phase, including the inputs, solution generation, evaluation, and filtering via the queue.

In the local search phase, only the solutions stored in the priority queue are refined. GRASPQ-FS explores the neighborhood of each solution by applying local modifications, typically replacing one feature at a time, and then re-evaluating their corresponding F1-Scores. If a better F1-Score is obtained, the solution is updated. This process continues iteratively for each queued solution until no further improvements are observed or the iteration limit is reached. The best solution found across all local searches is returned as the final selected feature subset. This phase is represented on the right side of Figure 5, emphasizing the iterative refinement and selection of the final output. The resulting subset contains exactly N features, as defined by the user at the beginning of the process [Quincozes *et al*., 2024].

5 Materials and Methods

In this section, we will detail the descriptions of the materials and methods used. This includes the datasets, preprocessing, classifier algorithms, experimentation scenarios, and experimental setup.

5.1 Data and Processing

In the experiments carried out, we used three different datasets, each extracted from a specific environment within the transmission layer, as summarized in Table 2.

Table 2. Summary of Datasets Used in the Study.

Dataset	Context	Features	Attacks Considered
X-CANIDS	Vehicular	688	Fabrication, Masquerade
CICIDS	Computer Networks	78	DDoS, PortScan
CICIoMT	Healthcare	45	Recon, Spoofing

The first dataset used was X-CANIDS [Jeong *et al*., 2024], collected from a Hyundai LF Sonata 2017 e-VGT vehicle, representing a realistic vehicle environment. This dataset is derived from the CAN network and contains 688 features extracted from the raw CAN frame structure, including message identifiers, alive counters, checksums, and payload fields. In other words, the captured CAN messages were deserialized to expose low-level communication attributes exchanged between Electronic Control Units (ECUs). The dataset comprises legitimate communications as well as malicious traffic, covering attacks such as fuzzing, fabrication, suspension, masquerade and replay.

For this work, fabrication and masquerade attacks were selected, as they present more subtle characteristics and represent significant challenges for automated detection. These attacks target the integrity and authenticity of in-vehicle communication over the CAN bus, directly exploiting its inherent lack of source authentication. Successfully identifying these manipulation and impersonation attempts is paramount for ensuring the safety and reliability of vehicular CPSs within a smart city context.

The second dataset used was CICIDS [Sharafaldin *et al*., 2018], originating from a network environment with 78 features simulating real conditions. This dataset consists of labeled network traffic, containing detailed information such as packet headers, flow-level features (e.g., flow duration, packet length statistics), source and destination IPs, ports, protocols, and attack types. It encompasses various attacks, such as botnet, brute force, DoS, and DDoS. For this study, we specifically used DDoS and PortScan attacks, which are highly relevant for network security analysis.

Finally, we employed CICIoMT [Dadkhah *et al*., 2024], a dataset representing the medical environment with information on 18 types of cyberattacks targeting 40 IoMT de-

vices. This dataset comprises 45 features and covers widely used protocols in healthcare devices, such as Wi-Fi, MQTT, and Bluetooth. The features include both packet-level and protocol-specific fields, such as signal strength, transmission rates, connection durations, and session metadata. For the present work, we specifically selected "reconnaissance" and "spoofing" attacks. These attack types were chosen because they represent foundational security challenges in diverse IoMT environments typical of smart healthcare. Reconnaissance is a crucial precursor phase for targeted attacks, while spoofing directly undermines device/user identity and access control mechanisms. Detecting these activities is vital for safeguarding sensitive patient data and ensuring the trustworthy operation of interconnected medical devices.

Our choice of these three datasets (including CICIoMT) was motivated by the search for a comprehensive representation of various scenarios present in a smart city. This approach goes beyond an analysis focused solely on sensors, contemplating other relevant dimensions of the urban environment.

We preprocessed all datasets using a standardized approach. We applied the `LabelEncoder` method from the `scikit-learn` library to all necessary columns, ensuring consistent encoding of categorical variables. In the CICIDS and CICIoMT datasets, we removed samples with null values to maintain data integrity. To ensure reproducibility, we split the data into 80% for training and 20% for testing, setting the `random_state` parameter to 42.

5.2 Classifiers

In the classification stage of our experiments, we selected XGBoost [Chen and Guestrin, 2016], LightGBM [Ke *et al*., 2017], and HistGradientBoosting (HistGradient) [Guryanov, 2019]. Our choice was based on their widespread adoption in intrusion detection tasks, their ability to handle high-dimensional and imbalanced data, and their distinct optimization strategies, which allow us to compare performance, speed, and stability across realistic scenarios. Although all these algorithms are rooted in decision tree learning, they follow different approaches, which we discuss below.

XGBoost implements the gradient boosting method, where trees are iteratively added to correct the residual errors generated by previous trees. Tree splits are performed with the objective of maximizing the reduction of the loss function at each node, taking into account information gains weighted by feature importance [Chen and Guestrin, 2016].

LightGBM, also based on gradient boosting, distinguishes itself by being faster and more memory-efficient. It employs a technique called Leaf-to-Leaf Growth, which prioritizes the expansion of nodes with the highest residual loss, rather than performing simultaneous splits on all leaves, as in XGBoost. This strategy often results in more asymmetric and deeper trees, optimizing overall performance [Ke *et al*., 2017].

Finally, HistGradient utilizes an approach based on the discretization of continuous feature values into discrete intervals (bins), creating histograms to accelerate and simplify split calculations. Unlike the preceding methods, its trees grow in a more controlled manner, typically with a fixed depth, which helps in avoiding excessive splits and contributes to greater model stability [Guryanov, 2019].

These methodological distinctions highlight the particularities and inherent benefits of each algorithm during the classification stage. We detail the discrepancies between these methods in Section 6, where we present and analyze the results obtained for each algorithm, also considering the different feature selection methods explored.

5.3 Experimentation Scenarios

We used the IWSS (Section 4.1), IWSHAP (Section 4.2), SHAP Ranking (Section 4.3), and GRASPQ-FS (Section 4.4) algorithms as reference methods for our experiments. These algorithms served as the primary comparative basis; however, we also explored combinations among them to form mixed methods. This approach aimed to perform a comprehensive analysis of their capabilities and interactions.

In addition to executing experiments within each of the domains considered in this work using each reference algorithm individually, we implemented variants that integrate their results to evaluate different approaches.

The following provides a summary of the combined algorithms used to form these new hybrid methods.

- **I-G-FS (IWSS GRASPQ Feature Selection):** This method performs a preliminary feature reduction using the IWSS algorithm, the result of which is subsequently processed by the GRASPQ-FS algorithm.
- **S-G-FS (SHAP GRASPQ Feature Selection):** In this case, the output of the SHAP Ranking is used as input for the GRASPQ-FS algorithm, producing a distinct final feature subset.
- **IS-G-FS (IWSHAP GRASPQ Feature Selection):** Similarly, this dataset is produced by feeding the set obtained by the IWSHAP algorithm into the GRASPQ-FS algorithm.
- **F-G-FS (Full GRASPQ Feature Selection):** In this method, the GRASPQ-FS algorithm is applied directly to the complete set of features of each dataset without preliminary feature reduction.

Additionally, throughout this work, we use the term `FULL` to represent a complete dataset without the application of feature selection techniques. This term consistently refers to one of the X-CANIDS, CICIDS, or CICIoMT datasets.

Regarding the specific configurations adopted for the GRASPQ-FS algorithm and its combinations with other methods, we conducted all experiments using the parameters presented in Table 3.

Table 3. Parameter values, based on Quincozes *et al*. [2024].

Parameter	Value
Number of Local Search Rounds	100
Number of Construction Rounds	100
Restricted Candidate List	30
Final Subset Length	5

5.4 Experimental Setup

We conducted the experiments on a computing platform that features mid-range hardware, specifically a tenth-generation Intel® Core™ i5 processor, 48 GB of RAM, and the Linux Ubuntu 24.04 LTS operating system.

6 Results

To perform a thorough evaluation of the effectiveness and efficiency of the algorithms and classifiers used in this study, we consider two primary factors: performance metrics and feature reduction capacity. We emphasize that a model achieving the highest performance metrics does not necessarily qualify as the optimal choice if its feature reduction capabilities are inadequate. Conversely, a model capable of reducing 99% of the features is of limited utility if its subsequent performance is sub-par. We discuss the results of this evaluation in the following sections.

6.1 Performance Evaluation

In Table 4, we present the F1-Scores recorded for each scenario, with the highest metric in each row highlighted in bold. Given the extensive number of scenarios we analyzed, we focus on the F1-Score as a comprehensive and robust indicator. This metric effectively encapsulates the balance between precision and recall, thereby providing a reliable measure of overall model performance. The results presented are based on an 80% training and 20% testing split of the dataset.

Before discussing the effect of each feature selection method, we can extract two general observations from our experiments: i) behavioral fluctuation across different datasets and attack types, and ii) performance variation from the classifier's perspective. In particular, performance (F1-Score) varies significantly between different datasets and attack types. For CICIDS (DDoS and PortScan), the classifiers achieve nearly perfect scores (99.99% or 100%) across all feature selection methods. X-CANIDS and CICIoMT datasets exhibit greater variability, indicating the impact of feature selection methods. However, from the classifiers' perspective, XGBoost consistently achieves high F1-Scores, especially in the X-CANIDS and CICIDS datasets. LightGBM also performs competitively, particularly in CICIDS and CICIoMT. HistGradient tends to have lower performance in some cases, particularly for the X-CANIDS dataset.

Once preliminary observations are discussed, an effective way to interpret the results presented in Table 4 is by focusing on the columns corresponding to each feature selection algorithm. In general, we can visualize the following key observations: IWSS and SHAP-based methods (SHAP, IWSHAP) often perform well, sometimes better than the FULL feature set. However, employing GRASPQ-FS as an attempt to enhance the feature subsets selected by SHAP (as illustrated by the S-G-FS results) generally leads to lower F1-Scores, suggesting that it may not be as effective as the other evaluated methods. An exception is observed for the CICIDS dataset, where the approach achieves comparatively better results, as discussed previously. Finally, F-G-FS and IS-G-FS perform inconsistently, excelling in some cases while performing poorly in others.

The poorest performance metrics were observed in S-G-FS and F-G-FS on the largest dataset (X-CANIDS). This indicates that both the SHAP and FULL methods are unsuitable for serving as an initial algorithm to compose a shortlist of candidate solutions for GRASPQ-FS. In turn, SHAP demonstrates strong performance when applied independently, emerging as the top-performing method in 14 out of 18 cases.

In summary, the results from the F1-Score perspective highlight that XGBoost and LightGBM are the best performing classifiers overall. Feature selection can significantly improve performance, through a considerable reduction of features, which directly impacts computational performance, and considerable increases in F1-Score, especially on challenging datasets such as X-CANIDS and CICIoMT. For simpler cases (such as CICIDS), feature selection has minimal impact, since the classifiers already achieve near-perfect performance. The IWSS and SHAP-based methods tend to provide the most significant benefit from feature selection, although they were not suitable as construction phase methods for GRASPQ-FS. Furthermore, the results presented in Table 4 indicate that the SHAP classification approach stands out as the most consistent strategy, particularly in scenarios that do not involve traditional networks. Although hybrid techniques were explored, the combination of different methods did not produce substantial performance improvements.

It is crucial to emphasize that these conclusions are drawn solely from performance metrics. Feature reduction is specifically analyzed in Section 6.2, where GRASPQ-FS demonstrates outstanding performance. Based on the experiments conducted so far, the ranking-based SHAP approach emerges as the most effective method to balance performance and simplicity in feature selection within the evaluated smart city scenarios. However, we must carefully consider the trade-off between the performance metrics discussed in this section and the dataset reduction capability (discussed in the next section).

6.2 Feature Reduction

In this section, we evaluate the feature reduction capabilities of the algorithms examined in this study. The results, presented in Table 5, include the total number of selected features and the corresponding reduction percentages.

The columns "IWSS" and "I-G-FS" in Table 5 present the feature reduction achieved by the IWSS algorithm and the additional reduction obtained by applying the GRASPQ-FS algorithm to the IWSS-reduced datasets, respectively. Likewise, "SHAP" and "S-G-FS", as well as "IWSHAP" and "IS-G-FS", indicate the percentage of feature reduction achieved by the SHAP Ranking and IWSHAP methods individually, along with the additional reduction obtained by applying GRASPQ-FS to the datasets initially reduced by these methods. Since GRASPQ-FS generates a subset of features with a fixed length predefined (set to 5 features by default [Quincozes *et al.*, 2024]), it consistently produces solutions that adhere to this constraint.

Based on the performance results discussed in Section 6.1,

Table 4. Performance results: F1-Score (%) achieved by each method.

Dataset	Attack	Classifier	IWSS	I-G-FS	SHAP	S-G-FS	IWSHAP	IS-G-FS	F-G-FS	FULL
XCANIDS (688)	Fabrication	XGBoost	93.15	83.16	**93.40**	88.52	72.86	72.86	72.86	91.96
		LightGBM	**90.84**	87.76	90.65	72.68	87.92	84.47	72.86	89.77
		HistGradient	88.41	84.75	**90.82**	72.68	87.58	82.17	72.86	89.96
	Masquerade	XGBoost	95.63	85.77	**96.72**	57.03	96.37	92.25	52.50	95.60
		LightGBM	91.34	84.33	93.34	58.60	**93.87**	88.49	59.86	92.65
		HistGradient	55.04	62.02	**93.97**	57.66	93.19	85.01	54.99	92.05
CICIDS (78)	DDoS	XGBoost	**100**	99.99	99.99	99.99	99.99	99.99	99.99	99.99
		LightGBM	99.99	**100**	99.99	99.99	99.99	99.99	**100**	99.99
		HistGradient	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	99.98	**99.99**
	PortScan	XGBoost	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**
		LightGBM	**99.99**	**99.99**	**99.99**	99.98	**99.99**	**99.99**	**99.99**	**99.99**
		HistGradient	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**
CICIoMT (45)	Recon	XGBoost	99.49	99.22	**99.51**	99.04	99.48	99.16	99.16	99.47
		LightGBM	99.43	99.03	**99.44**	98.86	99.42	99.18	99.09	99.43
		HistGradient	99.42	99.04	**99.44**	99.08	99.42	99.07	99.02	99.41
	Spoofing	XGBoost	93.83	90.70	**93.95**	90.97	93.80	91.52	90.95	93.66
		LightGBM	93.26	91.63	**93.43**	90.46	93.31	91.65	90.58	93.35
		HistGradient	93.24	91.32	**93.76**	90.87	93.37	91.42	91.14	93.37

Table 5. Feature Reduction Results (%) as a metric to assess the effectiveness of different approaches in minimizing the feature set.

Dataset	Attack	Classifier	IWSS	I-G-FS	SHAP	S-G-FS	IWSHAP	IS-G-FS	F-G-FS
XCANIDS (688)	Fabrication	XGBoost	95.06	99.27	94.33	99.27	**99.71**	**99.71**	99.27
		LightGBM	96.51	**99.27**	91.42	**99.27**	97.97	**99.27**	**99.27**
		HistGradient	95.20	**99.27**	80.52	**99.27**	95.20	**99.27**	**99.27**
	Masquerade	XGBoost	95.49	**99.27**	93.02	**99.27**	97.38	**99.27**	**99.27**
		LightGBM	95.20	**99.27**	93.02	**99.27**	97.24	**99.27**	**99.27**
		HistGradient	98.11	**99.27**	49.85	**99.27**	96.80	**99.27**	**99.27**
CICIDS (78)	DDoS	XGBoost	85.90	**93.59**	87.18	**93.59**	85.90	**93.59**	**93.59**
		LightGBM	87.18	**93.59**	88.46	**93.59**	84.62	**93.59**	**93.59**
		HistGradient	87.18	**93.59**	85.90	**93.59**	84.62	**93.59**	**93.59**
	PortScan	XGBoost	83.33	**93.59**	61.54	**93.59**	92.31	92.31	**93.59**
		LightGBM	87.18	**93.59**	87.18	**93.59**	84.62	**93.59**	**93.59**
		HistGradient	89.74	89.74	78.21	**93.59**	88.46	**93.59**	**93.59**
CICIoMT (45)	Recon	XGBoost	55.56	**88.89**	28.89	**88.89**	68.89	**88.89**	**88.89**
		LightGBM	46.67	**88.89**	26.67	**88.89**	68.89	**88.89**	**88.89**
		HistGradient	42.22	**88.89**	0.00	**88.89**	64.44	**88.89**	**88.89**
	Spoofing	XGBoost	53.33	**88.89**	33.33	**88.89**	66.67	**88.89**	**88.89**
		LightGBM	55.56	**88.89**	28.89	**88.89**	62.22	**88.89**	**88.89**
		HistGradient	53.33	**88.89**	0.00	**88.89**	68.89	**88.89**	**88.89**

it is evident that DDoS and PortScan attacks are easily detected, regardless of the feature selection method employed, or even in the absence of any selection method. In such cases where detection performance cannot be significantly improved, feature reduction can still streamline the analysis process by minimizing the amount of required information. A smaller set of features leads to faster and more efficient evaluation. In this context, all GRASP-based methods (*i.e.,* I-G-FS, S-G-FS, IS-G-FS, and F-G-FS) successfully reduced the total number of features by 93.59% for most classifier algorithms without compromising their detection performance. In particular, XGBoost demonstrates outstanding performance, achieving a perfect 100% F1-Score to detect DDoS attacks within the CICIDS dataset while reducing 93.59% of the features by I-G-FS.

Moreover, other key insights emerge from our analysis of Table 5. In particular, IWSS demonstrates its strength in handling larger datasets, where it achieves significant feature reductions. However, its efficiency diminishes with smaller datasets, such as the CICIoMT dataset, where the reduction is less pronounced. In contrast, the performance of the GRASPQ-FS variants is particularly noteworthy with smaller datasets, where other algorithms struggle. Therefore, although GRASPQ-FS did not achieve the highest F1-Score, as observed in the previous section, it proved to be the most effective method to reduce the number of features. This makes it a suitable choice for feature selection in computationally constrained smart city scenarios, where feature reduction may be mandatory.

Despite its strengths, GRASPQ-FS-based algorithms have inherent limitations, particularly their inability to operate on datasets with a reduced number of features in the RCL. This limitation became evident in three instances during the experiments: twice for the IS-G-FS method and once for the I-G-FS algorithm.

The first instance of this behavior occurred in the X-

CANIDS dataset during a fabrication attack, where the IWSHAP method, when combined with the XGBoost classifier, generated a feature subset containing only two features. The second instance was observed in the CICIDS dataset during a portscan attack with XGBoost, where IWSHAP produced a final subset of six features. The third case also happened within the CICIDS dataset, during a portscan attack using the HistGradient classifier. In this scenario, the IWSS method resulted in a feature set of eight, which proved insufficient for the GRASPQ-FS phase within the I-G-FS method to execute properly. This insufficiency is explained by the experimental configuration, in which GRASPQ-FS, although set to build an RCL of size 30, receives a dynamic feature input in the I-G-FS scenario. In this specific case, the number of available features was smaller than 30, which led to the insufficiency of GRASPQ-FS.

In the three specific cases where the GRASPQ-FS phase could not be executed, the reduction percentage for methods combining a baseline algorithm with GRASPQ-FS (such as IS-G-FS and I-G-FS) remained identical to that achieved by the baseline method alone. This indicates that, in certain configurations, the algorithm's local search phase may not provide additional benefits beyond the initial reduction process.

For the SHAP experiments, we conducted an extensive analysis to determine the optimal cut-off threshold for feature selection. The number of features retained at the top of the ranking varied according to the classifier and dataset used. Under this approach, SHAP generally demonstrated slightly lower reduction capability than IWSS, especially in smaller datasets. In some cases, SHAP failed to remove any features; instead, it reordered their

Furthermore, Table 5 highlights the synergistic potential of combining the IWSS and SHAP algorithms. This hybrid approach consistently outperforms its individual components, even in challenging scenarios such as the CICIoMT dataset, where both IWSS and SHAP individually performed worse in terms of feature reduction. A particularly notable example is the fabrication attack evaluated with the XGBoost classifier, in which the number of features was reduced from 688 to only two, representing a reduction of 99.71%. However, it is crucial to consider the corresponding F1-Score, which, at 72.86%, is insufficient for practical applications. Therefore, considering only feature reduction in cases like this is neither fair nor helpful.

Finally, the standalone performance of the GRASPQ-FS algorithm, applied to complete datasets (F-G-FS), is detailed. The results confirm GRASPQ-FS's capability for aggressive feature reduction, achieving substantial reductions even for large datasets. However, its fixed output of five features introduces a limitation. For example, reducing a dataset with 688 features to only five can lead to suboptimal performance metrics. This is evident in the X-CANIDS dataset, where the masquerade attack yielded an F1-Score of approximately 50%, and the fabrication attack reached around 72%, despite a reduction of 99.27%.

Although the GRASPQ-FS algorithm offers flexibility in customizing the final number of selected features, we found that achieving an optimal balance often requires a trial-and-error approach, necessitating multiple iterations. This iterative process presents a disadvantage when compared to other algorithms that do not demand such fine-tuning, making them more straightforward and efficient in practical applications.

We present more information on GRASPQ-FS variants in Section 6.3, where we conduct a comprehensive analysis of feature reduction during its construction and local search phases.

6.3 GRASP-based Feature Reduction

In this section, we analyze the influence of the construction and local search phases in GRASP-based methods (*i.e.*, those ending with "-G-FS"), providing deeper insights into their effectiveness in feature selection.

In Figure 6, detailed information is presented about the GRASP-based algorithms, illustrating how their construction and local search phases perform in terms of feature reduction across different attacks, datasets, and classifier algorithms.

Since all feature selection methods in Figure 6 are based on GRASPQ-FS, the implementation used in this work enforces a fixed-size output, resulting in all solutions reaching the same final feature reduction. However, the key differences lie in the magnitude of reduction achieved in each phase, construction and local search. The I-G-FS and IS-G-FS methods, which are based on IWSS and IWSS with SHAP, tend to perform more aggressive reductions in the construction phase, consequently limiting the search space available for the local search phase. This pattern is particularly noticeable in datasets with fewer features, such as CICIoMT, where the initial reduction exceeds 60% in some cases.

Another trend is the behavior of the HistGradient model, which tends to retain more features during the construction phase. When combined with the S-G-FS technique, which employs SHAP-Ranking during this stage, the percentage of feature reduction in the construction phase is significantly lower, reaching approximately 50% in the case of the Masquerade attack within the largest dataset (X-CANIDS). An even smaller reduction is observed for the Spoofing attack within the smallest dataset (CICIoMT), where no feature reduction occurred during the construction phase, with all reduction taking place in the local search phase. In particular, CICIoMT generally exhibited a lower impact on feature reduction in the construction phase, except for the case of HistGradient processing the Recon attack, where an anomalous behavior led to a significant feature reduction in the SHAP-Ranking-based construction phase. This suggests that the Recon attack can be effectively described with a smaller set of features, according to the observed SHAP values.

In general, the IS-G-FS and I-G-FS methods, which leverage IWSS, tend to produce more aggressive reductions in the construction phase. This, in turn, reduces the search space and restricts the alternatives available for feature reduction in the local search phase. The size of the dataset also plays a crucial role in the reduction pattern: datasets with fewer features, such as CICIoMT, experience a more pronounced impact in the construction phase. Meanwhile, the S-G-FS method exhibits a more gradual reduction, particularly when using HistGradient, due to its tendency to retain a higher number of features early on.

These findings highlight the importance of choosing an

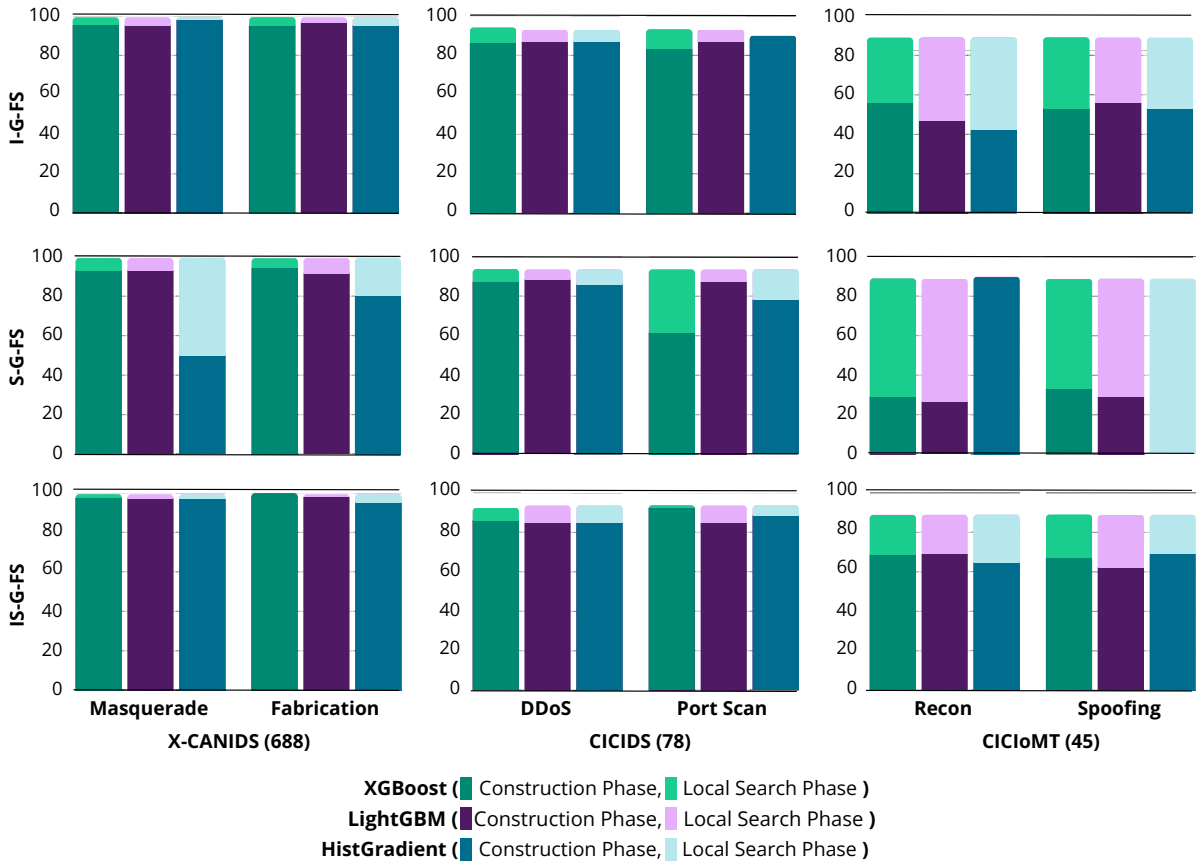


Figure 6. Feature reduction (%) by GRASP-based variants with XGBoost, LightGBM, and HistGradient on different datasets and attack types.

appropriate feature selection strategy based on both the characteristics of the dataset and the types of attacks analyzed.

6.4 Impact of Selected Features on IDS

This section provides an overview of how the most influential features, identified by hybrid and XAI-based methods, correspond to the expected behavior of different attack types. While Section 6 focuses on performance metrics and feature reduction, here we highlight representative examples that illustrate the alignment between selected features and the operational characteristics of the evaluated attacks across different domains.

In the vehicular CPS scenario (X-CANIDS dataset [Jeong *et al.*, 2024]), the features selected for detecting Fabrication and Masquerade attacks reveal clear alignment with the internal logic of CAN-based threats. For **Fabrication** attacks, highly influential features such as `2B0_SAS_Angle`, `2B0_SAS_Stat`, and `50C_CF_Clu DTE` indicate manipulations targeting steering stability and fuel estimation, consistent with the tactic of injecting crafted messages that mimic legitimate traffic but distort critical control information. The presence of `5B0_CF_Clu_Odometer` as a recurrent feature across both attacks suggests attackers aim to falsify metrics displayed to drivers. In **Masquerade** scenarios, although less extensively discussed in the dataset, similar dependencies on identifiers and payload consistency are expected, reinforcing the relevance of features tied to message structure and ECU-specific semantics. Altogether, these findings support the notion that meaningful feature selection in CAN networks must capture not only statistical variance, but also reflect the adversarial strategies aimed at specific vehicular subsystems.

In the **IoMT domain** (CICIoMT), attacks such as Recon and Spoofing target device identification and communication metadata. The feature selection process consistently prioritized attributes related to signal strength, transmission rate, and session identifiers. These variables are essential for identifying anomalies in access patterns or unexpected communication flows, supporting detection mechanisms that prevent unauthorized access or data manipulation in sensitive medical contexts.

For **Reconnaissance** attacks, the consistently most important features include flag counts (`fin_count`, `rst_count`, `syn_count`), packet rates (`Rate`), packet sizes (`Tot size`, `Max`, `Tot sum`), inter-arrival time (`IAT`), and header information (`Header_Length`). The relevance of these characteristics stems from the fact that reconnaissance attacks generally involve probing ports and services, generating traffic with atypical characteristics, such as incomplete connections or elevated packet rates for short periods. The presence of connection flags and atypical flow metrics are direct indicators of scanning activities. Furthermore, `Protocol Type` and the identification of the use of specific protocols (*e.g.,* `UDP`, `HTTP`, `HTTPS`, `TCP`, `ICMP`, `Telnet`, `SMTP`) also prove important, suggesting that the IDS can detect the exploitation of unusual services or protocols in medical devices during the reconnaissance phase.

In **Spoofing** attacks, where identity falsification is the core

of the threat, features such as packet rates (`Rate`, `Srate`), inter-arrival time (`IAT`), and packet size statistics (`Max`, `Min`, `AVG`, `Tot size`) are crucial. This is because a spoofed device or user may exhibit traffic patterns that deviate significantly from legitimate behavior, such as sending data at irregular rates or intervals, or with packet sizes that are not typical of the emulated device. The relevance of `Protocol Type` and flags (*e.g.,* `psh_flag_number`, `ack_flag_number`, `rst_count`, `syn_count`) is also observed, as an attacker might attempt to use unexpected packet sequences or protocols to maintain falsification or deceive the system. Detecting these behavioral anomalies is vital for communication integrity and patient data privacy.

For the computer network scenario using the CICIDS dataset Sharafaldin *et al*. [2018], our evaluation of DDoS and PortScan attacks relied predominantly on statistical and flow-based features. Examples include `Total Fwd Packets`, `Flow Duration`, and `Flow Packets/s`. These attributes proved instrumental in capturing volumetric anomalies and scanning behaviors, which are characteristic of reconnaissance and service disruption attempts. The consistent appearance of these features among the top-ranked across various selection strategies reinforces their critical importance in baseline network protection.

For **DDoS** attacks, the consistently most relevant features include `Total Length of Fwd Packets`, `Total Length of Bwd Packets`, `Total Fwd Packets`, `Total Backward Packets`, `Flow Duration`, and `Destination Port`. The high relevance of these characteristics is justified by the very volumetric nature of DDoS, which aims to overload network resources. The detection of abnormal data volumes (*e.g.,* `Total Length of Fwd Packets`, `Total Fwd Packets`) and the identification of multiple connection attempts to a specific `Destination Port` are essential for identifying the attack. Furthermore, features related to connection flags (*e.g.,* `SYN Flag Count`, `FIN Flag Count`, `ACK Flag Count`) and initial window sizes (*e.g.,* `Init_Win_bytes_forward`, `Init_Win_bytes_backward`) also prove important, as they can indicate abnormal handshake patterns or atypical connection terminations in an overload attack.

In the case of **Port Scan** attacks, `Destination Port` emerges as the most critical feature, which is expected, given that the attack consists of probing multiple ports to identify vulnerabilities. Additionally, features such as `Flow Duration`, `Packet Length Mean`, `Bwd Packet Length Max`, `Flow IAT Std`, `Fwd IAT Mean`, `Flow Bytes/s`, `Flow Packets/s`, `Flow IAT Min`, `Fwd IAT Total`, and `PSH Flag Count` are also highly relevant. These features allow the IDS to identify unusual scanning patterns, such as many short connections to different ports, atypical packet sizes, or irregular inter-arrival times that do not result in normal communication, being crucial for effective network reconnaissance detection.

The correspondence observed between the selected features and the operational patterns of each attack indicates that these methods can reveal how intrusions manifest at the data level. Thus, beyond enhancing performance metrics, feature selection also supports the interpretation of adversarial behavior, reinforcing its role as a strategic component in the design of transparent and security-aware IDS models.

6.5 Discussion

After evaluating the performance of these algorithms in various scenarios, considering both the F1-Score and the capability to reduce features, the nuances and particularities of each technique become evident. By combining insights from Table 4 (F1-Score performance) and Table 5 (Feature Reduction effectiveness), we can analyze the key trade-offs between maintaining classification performance and reducing the number of features.

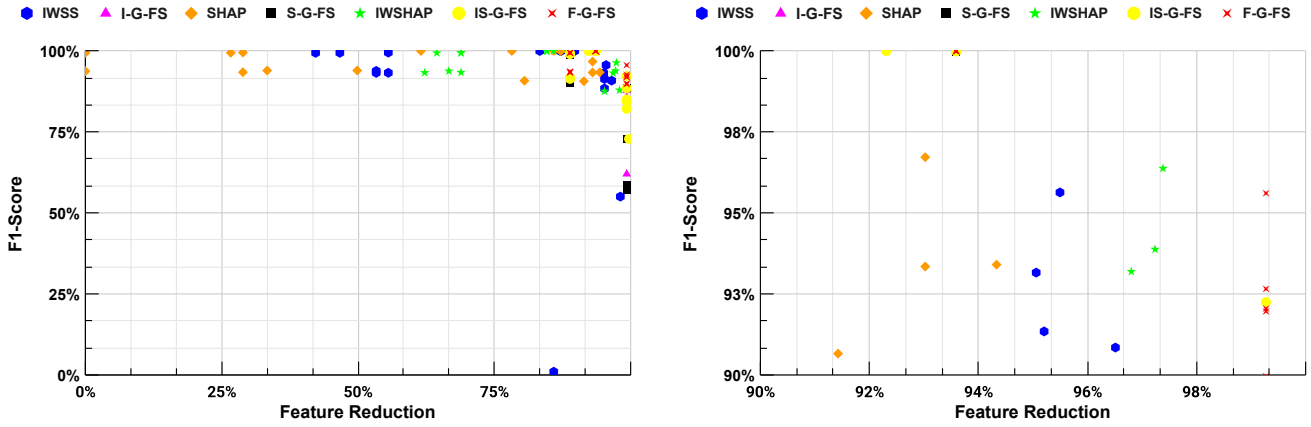
The scatter plots in Figure 7 visually highlight the trade-offs between classification performance (F1-Score) and feature reduction in different feature selection methods. From the overall distribution in Figure 7a, it is evident that most of the methods maintain high F1-Scores, even with a substantial reduction in features. However, there are noticeable variations between different techniques, emphasizing that the effectiveness of feature selection depends on the dataset and the method applied.

One key observation is that SHAP-based methods, particularly SHAP and IWSHAP, provide a strong balance between performance and feature reduction. These methods appear in the upper right region of the plots, indicating that they can remove a significant number of features while still preserving high classification accuracy. This suggests that SHAP-based approaches are well-suited for scenarios where both accuracy and computational efficiency are priorities.

It is important to emphasize that feature reduction has a direct and measurable impact on system performance. Beyond lowering computational complexity, our results demonstrate that feature selection, when performed with suitable algorithms such as SHAP or IWSHAP, can also improve classification accuracy. For example, in the X-CANIDS dataset, IWSHAP reduced the feature set by more than 99% while achieving an F1-Score above 93%, outperforming the model trained with all features. These findings confirm that feature reduction is not merely an optimization for efficiency but also a factor that directly contributes to model effectiveness in complex smart city scenarios.

In contrast, GRASPQ-FS-based methods (IS-G-FS and F-G-FS) demonstrate extreme feature reduction capabilities, often reaching reductions above 88.89%. However, as seen in the CICIDS dataset, if the initial feature selection step leaves too few features, the GRASPQ-FS phase may be unable to proceed, which leads to no additional reduction. This highlights a potential limitation of these methods in cases where the selection of the baseline features is overly aggressive.

It is also important to consider feature selection not as a static preprocessing step, but as a continuous and adaptive process that can operate in parallel with the deployed IDS. In dynamic environments such as smart cities, network behavior and attack patterns may evolve over time, potentially reintroducing the relevance of previously discarded features. Therefore, maintaining the ability to revisit and reconstruct feature sets in response to traffic shifts is essential to preserving detection effectiveness. This perspective aligns with recent research efforts that propose continuous and distributed feature selection mechanisms, such as GDLS-FS [Silva *et al*.,



(a) Overall relation between F1-Score and Feature Reduction. **(b)** Best trade-off relation between F1-Score and Feature Reduction (Zoomed-in view).

Figure 7. Comparison of F1-Score and Feature Reduction across different feature selection methods.

2023], which integrates GRASP-based heuristics with local search to support adaptive feature selection in real-time intrusion detection scenarios.

Furthermore, it is worth highlighting that although the methods present a considerable computational cost during their process, especially SHAP Ranking, the computational performance is improved with the result of these processes, through the reduction of characteristics, which makes the processes computationally lighter, leading to better performance [Scherer *et al*., 2024a].

7 Conclusion

This study conducted a comprehensive evaluation of feature selection techniques for IDSs in smart city environments, addressing the critical trade-offs between classification performance and feature reduction. We analyze a diverse range of methods, including XAI-driven techniques (SHAP, IWSHAP), traditional wrapper techniques (IWSS), and hybrid metaheuristic-based techniques (GRASPQ-FS and its variations).

Our results revealed that SHAP-driven methods (SHAP, IWSHAP) consistently achieved high classification performance while maintaining substantial feature reduction, positioning them as a robust choice for balancing accuracy and efficiency. In contrast, GRASPQ-FS demonstrated remarkable feature reduction, often exceeding 90%, but its performance was highly dependent on the initial feature selection phase. In certain instances, when the initial reduction was excessively aggressive, GRASPQ-FS was unable to effectively refine the feature set, limiting its overall effectiveness.

Hybrid approaches that combined GRASPQ-FS with SHAP or IWSS did not produce significant improvements in classification performance, suggesting that these methods are best suited for scenarios requiring extreme feature reduction rather than enhancement of detection performance. Moreover, we observed that the size of the dataset and the type of attack significantly influenced the effectiveness of feature selection techniques, reinforcing the need for adaptive strategies tailored to specific IDS use cases.

In future work, we plan to explore adaptive feature selection frameworks that dynamically adjust selection strategies based on dataset characteristics, as well as investigate the integration of additional machine learning models to enhance IDSs' robustness. Our findings provide valuable information for optimizing IDSs deployment in smart city environments, where efficiency, interpretability, and robust cybersecurity are paramount.

Declarations

Authors' Contributions

F.H.S., F.N.D., and M.M.C. contributed to the conceptualization, implementation, and execution of the experiments. S.E.Q. and D.K. provided crucial methodological insights and supervised the research process. V.E.Q. contributed to the development of the theoretical framework and the critical review of the manuscript. All authors actively participated in the discussion of the results, the drafting of the manuscript, and the final approval of the submitted version.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We express our sincere gratitude to the Programa de Desenvolvimento Acadêmico (PDA) of the Federal University of Pampa (UNIPAMPA) and the Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS) for their generous financial support through scholarships and infrastructure funding. The authors also thank the Laboratory of Advanced Studies in Computing (LEA) at UNIPAMPA for providing the essential computational resources that enabled this research. Finally, we are deeply grateful for the insightful discussions and constructive feedback from our colleagues and reviewers, which significantly contributed to the refinement and improvement of this work.

Funding

This research was supported by the Programa de Desenvolvimento Acadêmico (PDA) of the Federal University of Pampa (UNI-

PAMPA), which provided financial assistance through scholarships. Additionally, we acknowledge that this work was partially funded by the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) through grant agreements 24/2551-0001368-7 and 24/2551-0000726-1, which contributed to infrastructure and research development support. We express our gratitude for the financial aid that enabled the execution of this study.

Availability of data and materials

The IWSHAP feature selection tool, presented in Scherer *et al*. [2024b], is publicly available in the following repository: `https://github.com/sf24-iwshap/sf24-iwshap`.

The datasets analyzed during the current study are publicly available from their original sources (X-CANIDS Jeong *et al*. [2024], CICIDS Sharafaldin *et al*. [2018], CICIOMT Dadkhah *et al*. [2024]), although access or redistribution may be subject to their respective licenses or terms of use.

The specific code developed and used for the analyses presented in this study is available from the corresponding author upon reasonable request. Due to ethical and security considerations related to potentially sensitive configurations or dataset handling, direct public release of the analysis scripts is restricted. For further information, please contact the corresponding author.

References

Abbas, T., Khan, A. H., Kanwal, K., Daud, A., Irfan, M., Bukhari, A., and Alharbey, R. (2024). IoMT-Based Healthcare Systems: A Review. *Computer Systems Science & Engineering*, 48(4). DOI: 10.32604/csse.2024.049026.

Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., and Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*, 32(1):e4150. DOI: 10.1002/ett.4150.

Arreche, O., Guntur, T., and Abdallah, M. (2024). Xai-based feature selection for improved network intrusion detection systems. DOI: 10.48550/arxiv.2410.10050.

Ashibani, Y. and Mahmoud, Q. H. (2017). Cyber physical systems security: Analysis, challenges and solutions. *Computers & Security*, 68:81–97. DOI: 10.1016/j.cose.2017.04.005.

Awajan, A. (2023). A novel deep learning-based intrusion detection system for iot networks. *Computers*, 12(2). DOI: 10.3390/computers12020034.

Bakro, M., Kumar, R. R., Husain, M., Ashraf, Z., Ali, A., Yaqoob, S. I., Ahmed, M. N., and Parveen, N. (2024). Building a cloud-ids by hybrid bio-inspired feature selection algorithms along with random forest model. *IEEE Access*, 12:8846–8874. DOI: 10.1109/ACCESS.2024.3353055.

Bataineh, A. S., Zulkernine, M., Abusitta, A., and Halabi, T. (2024). Detecting poisoning attacks in collaborative idss of vehicular networks using xai and shapley value. *Journal on Autonomous Transportation Systems*, 2(3):1–21. DOI: 10.1145/3696462.

Bermejo, P., Gamez, J. A., and Puerta, J. M. (2009). Incremental wrapper-based subset selection with replacement: An advantageous alternative to sequential forward selection. In *2009 IEEE Symposium on Computational Intelligence and Data Mining*, pages 367–374. DOI: 10.1109/CIDM.2009.4938673.

Cao, L., Jiang, X., Zhao, Y., Wang, S., You, D., and Xu, X. (2020). A survey of network attacks on cyber-physical systems. *IEEE Access*, 8:44219–44227. DOI: 10.1109/ACCESS.2020.2977423.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28. 40th-year commemorative issue. DOI: 10.1016/j.compeleceng.2013.11.024.

Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/2939672.2939785.

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., and Stoddart, K. (2016). A review of cyber security risk assessment methods for scada systems. *Computers & Security*, 56:1–27. DOI: 10.1016/j.cose.2015.09.009.

Dadkhah, S., Neto, E. C. P., Ferreira, R., Molokwu, R. C., Sadeghi, S., and Ghorbani, A. A. (2024). Ciciomt2024: A benchmark dataset for multi-protocol security assessment in iomt. *Internet of Things*, 28:101351. DOI: 10.1016/j.iot.2024.101351.

Demertzi, V., Demertzis, S., and Demertzis, K. (2023). An overview of cyber threats, attacks and countermeasures on the primary domains of smart cities. *Applied Sciences*, 13(2). DOI: 10.3390/app13020790.

Diaba, S. Y., Shafie-khah, M., and Elmusrati, M. (2024). Cyber-physical attack and the future energy systems: A review. *Energy Reports*, 12:2914–2932. DOI: 10.1016/j.egyr.2024.08.060.

Din, I. U., Almogren, A., and Rodrigues, J. J. (2024). AIoT Integration in Autonomous Vehicles: Enhancing Road Cooperation and Traffic Management. *IEEE Internet of Things Journal*. DOI: 10.1109/JIOT.2024.3387927.

Divekar, A., Parekh, M., Savla, V., Mishra, R., and Shirole, M. (2018). Benchmarking datasets for anomaly-based network intrusion detection: KDD CUP 99 alternatives. *CoRR*, abs/1811.05372. DOI: 10.1109/CCCS.2018.8586840.

Dui, H., Zhang, S., Liu, M., Dong, X., and Bai, G. (2024). IoT-Enabled Real-Time Traffic Monitoring and Control Management for Intelligent Transportation Systems. *IEEE Internet of Things Journal*, 11(9):15842–15854. DOI: 10.1109/JIOT.2024.3351908.

E. L. Asry, C., Benchaji, I., Douzi, S., and E. L. Ouahidi, B. (2024). A robust intrusion detection system based on a shallow learning model and feature extraction techniques. *PLOS ONE*, 19(1):1–31. DOI: 10.1371/journal.pone.0295801.

Esseghir, M. A. (2010). Effective wrapper-filter hybridization through grasp schemata. Available at:`https://proceedings.mlr.press/v10/esseghir10a.html`.

Evancich, N. and Li, J. (2016). *Attacks on Industrial Control Systems*, pages 95–110. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-32125-7_6.

Faliagka, E., Christopoulou, E., Ringas, D., Politi, T., Kostis, N., Leonardos, D., Tranoris, C., Antonopoulos, C. P., Denazis, S., and Voros, N. (2024). Trends in digital twin framework architectures for smart cities: A case study in smart mobility. *Sensors*, 24(5):1665. DOI: 10.3390/s24051665.

Fang, Y., Yao, Y., Lin, X., Wang, J., and Zhai, H. (2024). A feature selection based on genetic algorithm for intrusion detection of industrial control systems. *Computers & Security*, 139:103675. DOI: 10.1016/j.cose.2023.103675.

Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., and Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *computers & security*, 28(1-2):18–28. DOI: 10.1016/j.cose.2008.08.003.

Guryanov, A. (2019). Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees. In van der Aalst, W. M. P., Batagelj, V., Ignatov, D. I., Khachay, M., Kuskova, V., Kutuzov, A., Kuznetsov, S. O., Lomazova, I. A., Loukachevitch, N., Napoli, A., Pardalos, P. M., Pelillo, M., Savchenko, A. V., and Tutubalina, E., editors, *Analysis of Images, Social Networks and Texts*, pages 39–50, Cham. Springer International Publishing. DOI: 10.1007/978-3-030-37334-4_4.

Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3(null):1157–1182. DOI: 10.5555/944919.944968.

Habibi Lashkari, A., Kaur, G., and Rahali, A. (2020). Didarknet: A contemporary approach to detect and characterize the darknet traffic using deep image learning. In *Proceedings of the 2020 10th International Conference on Communication and Network Security*, pages 1–13. DOI: 10.1145/3442520.3442521.

Javeed, D., Gao, T., Saeed, M. S., Kumar, P., Kumar, R., and Jolfaei, A. (2023). A softwarized intrusion detection system for iot-enabled smart healthcare system. *ACM Trans. Internet Technol.* Just Accepted. DOI: 10.1145/3634748.

Jeong, S., Lee, S., Lee, H., and Kim, H. K. (2024). X-CANIDS: Signal-aware explainable intrusion detection system for controller area network-based in-vehicle network. *IEEE Transactions on Vehicular Technology*, 73(3):3230–3246. DOI: 10.1109/TVT.2023.3327275.

Kato, T., Fukumoto, N., Sasaki, C., Tagami, A., and Nakao, A. (2024). Challenges of cps/iot network architecture in 6g era. *IEEE Access*. DOI: 10.1109/ACCESS.2024.3395363.

Kayan, H., Nunes, M., Rana, O., Burnap, P., and Perera, C. (2022). Cybersecurity of industrial cyber-physical systems: A review. *ACM Comput. Surv.*, 54(11s). DOI: 10.1145/3510410.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). Lightgbm: a highly efficient gradient boosting decision tree. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 3149–3157, Red Hook, NY, USA. Curran Associates Inc.. DOI: 10.5555/3294996.3295074.

Khani, P., Moeinaddini, E., Abnavi, N. D., and Shahraki, A. (2024). Explainable artificial intelligence for feature selection in network traffic classification: A comparative study. *Transactions on Emerging Telecommunications Technologies*, 35(4):e4970. DOI: 10.1002/ett.4970.

Kim, D.-j., NG, B. A., and Sarveshwaran, V. (2024). A novel split learning based consumer electronics network traffic anomaly detection framework for smart city environment. *IEEE Transactions on Consumer Electronics*. DOI: 10.1109/TCE.2024.3367330.

Kwon, D., Kim, H., Kim, J., Suh, S. C., Kim, I., and Kim, K. J. (2019). A survey of deep learning-based network anomaly detection. *Cluster Computing*, 22:949–961. DOI: 10.1007/s10586-017-1117-8.

Li, J., Othman, M. S., Chen, H., and Yusuf, L. M. (2024). Optimizing iot intrusion detection system: feature selection versus feature extraction in machine learning. *Journal of Big Data*, 11(1):36. DOI: 10.1186/s40537-024-00892-y.

Mahmoud, R., Yousuf, T., Aloul, F., and Zualkernan, I. (2015). Internet of things (iot) security: Current status, challenges and prospective measures. In *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, pages 336–341. DOI: 10.1109/ICITST.2015.7412116.

Marwedel, P. (2021). *Embedded System Design: Embedded Systems Foundations of Cyber-Physical Systems, and the Internet of Things*. Springer Nature. DOI: 10.1007/978-3-030-60910-8.

Mihailescu, M.-E., Mihai, D., Carabas, M., Komisarek, M., Pawlicki, M., Hołubowicz, W., and Kozik, R. (2021). The proposition and evaluation of the roedunet-simargl2021 network intrusion detection dataset. *Sensors*, 21(13). DOI: 10.3390/s21134319.

Mitchell, R. and Chen, I.-R. (2014). A survey of intrusion detection techniques for cyber-physical systems. *ACM Comput. Surv.*, 46(4). DOI: 10.1145/2542049.

Moradkhani, M., Amiri, A., Javaherian, M., and Safari, H. (2015). A hybrid algorithm for feature subset selection in high-dimensional datasets using fica and iwssr algorithm. *Applied Soft Comp.*, 35:123. DOI: 10.1016/j.asoc.2015.03.049.

Okada, S., Jmila, H., Akashi, K., Mitsunaga, T., Sekiya, Y., Takase, H., Blanc, G., and Nakamura, H. (2025). Xai-driven black-box adversarial attacks on network intrusion detectors. *International Journal of Information Security*, 24(3):1–15. DOI: 10.1007/s10207-025-01016-0.

Pesaramelli, R. S. and Sujatha, B. (2024). Principle correlated feature extraction using differential evolution for improved classification. In *AIP Conference Proceedings*, volume 2919. AIP Publishing. DOI: 10.1063/5.0184938.

Quincozes, S. E., Passos, D., Albuquerque, C., Ochi, L. S., and Mossé, D. (2020). GRASP-based feature selection for intrusion detection in cps perception layer. In *2020 4th Conference on Cloud and Internet of Things (CIoT)*, pages 41–48. DOI: 10.1109/CIoT50422.2020.9244207.

Quincozes, V. E., Quincozes, S. E., Albuquerque, C., Passos, D., and Mossé, D. (2024). Efficient feature selection for intrusion detection systems with priority queue-based grasp. In *2024 IEEE 13th International Conference on Cloud Networking (CloudNet)*, pages 1–8. DOI: 10.1109/CloudNet62863.2024.10815746.

Saber, O. and Mazri, T. (2021). Smart city security issues: The main attacks and countermeasures. *Int. Arch.*

Photogramm. Remote Sens. Spatial Inf. Sci., XLVI-4/W5-2021:465–472. DOI: 10.5194/isprs-archives-XLVI-4-W5-2021-465-2021.

Scherer, F., Dresch, F., Quincozes, S., Kreutz, D., and Quincozes, V. (2024a). IWSHAP: Um método de seleção incremental de características para redes can baseado em inteligência artificial explicável (XAI). In *Anais do XXIV Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais*, pages 351–366, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbseg.2024.241780.

Scherer, F., Dresch, F., Quincozes, S., Kreutz, D., and Quincozes, V. (2024b). Iwshap: Uma ferramenta para seleção incremental de características utilizando iwss e shap. In *Anais Estendidos do XXIV Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais*, pages 105–112, Porto Alegre, RS, Brasil. SBC. DOI: 10.5753/sbseg_estendido.2024.243376.

Shanbhag, A., Vincent, S., Gowda, S. B. B., Kumar, O. P., and Francis, S. A. J. (2024). Leveraging metaheuristics for feature selection with machine learning classification for malicious packet detection in computer networks. *IEEE Access*, 12:21745–21764. DOI: 10.1109/ACCESS.2024.3362246.

Sharafaldin, I., Lashkari, A. H., Ghorbani, A. A., *et al*. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSp*, 1:108–116. DOI: 10.5220/0006639801080116.

Silva, E. F., Naves, N., Quincozes, S. E., Quincozes, V. E., Kazienko, J. F., and Cheikhrouhou, O. (2023). GDLS-FS: Scaling feature selection for intrusion detection with grasp-fs and distributed local search. In *International conference on advanced information networking and applications*, pages 199–210. Springer. DOI: 10.1007/978-3-031-28451-9$_1$8.

Sivamohan, S. and Sridhar, S. (2023). An optimized model for network intrusion detection systems in industry 4.0 using xai based bi-lstm framework. *Neural Computing and Applications*, 35(15):11459–11475. DOI: 10.1007/s00521-023-08319-0.

Stutz, D., de Assis, J. T., Laghari, A. A., Khan, A. A., Deshpande, A., Kulkarni, D., Terziev, A., de Jesus, M. A., and Grata, E. G. (2024). Cyber threat detection and mitigation using artificial intelligence–a cyber-physical perspective. *Applying Artificial Intelligence in Cybersecurity Analytics and Cyber Threat Detection*, pages 107–133. DOI: 10.1002/9781394196470.ch7.

Su, R. (2018). Supervisor synthesis to thwart cyber attack with bounded sensor reading alterations. *Automatica*, 94:35–44. DOI: 10.1016/j.automatica.2018.04.006.

Tavallaee, M., Bagheri, E., Lu, W., and Ghorbani, A. A. (2009). A detailed analysis of the kdd cup 99 data set. In *2009 IEEE symposium on computational intelligence for security and defense applications*, pages 1–6. Ieee. DOI: 10.1109/CISDA.2009.5356528.

Turukmane, A. V. and Devendiran, R. (2024). M-multisvm: An efficient feature selection assisted network intrusion detection system using machine learning. *Computers & Security*, 137:103587. DOI: 10.1016/j.cose.2023.103587.