Hi-Fi Flow: Real-Time High-Granularity Flow Feature Extraction for Robust Network Monitoring

Shengping Bi, Lang Zhou, Tao Wang, and Tao Hou
Department of Computer Science and Engineering
University of North Texas
Denton, TX, USA
{shengpingbi, langzhou}@my.unt.edu, {tao.wang2, tao.hou}@unt.edu

***Abstract*— Modern networks exhibit rapidly expanding traffic volume, diversity, and dynamism, which hinder accurate flow reconstruction and real-time monitoring. Existing flow feature extraction tools primarily rely on static 5-tuple attributes and post-completion processing, resulting in fragmented flow records and delayed detection under evasive behaviors such as IP spoofing, port rotation, and multi-session tunneling. We present Hi-Fi Flow, a real-time, high-granularity flow identification and feature extraction tool that jointly models spatial, temporal, and structural correlations among packets. Hi-Fi Flow employs a dual self-attention encoder to capture complementary endpoint-level and behavior-level dependencies for fine-grained packet-to-flow association. To reduce labeling overhead, it further adopts a few-shot triplet learning strategy that learns discriminative packet-flow similarity patterns with limited labeled data. Extensive evaluations demonstrate near-perfect flow identification accuracy, reaching 99∼100% with only 50 labeled samples per attack, and show that Hi-Fi Flow features consistently outperform legacy flowmeters in downstream attack detection tasks. The proposed framework provides an efficient and robust foundation for next-generation network monitoring and threat detection.**

I. Introduction

The expansion of cloud computing, IoT deployments, and emerging AI workloads has driven an explosive increase in network traffic volume, diversity, and dynamism [1], [2]. As traffic now traverses heterogeneous devices, dynamic routing paths, and heavily encrypted channels, it creates considerable challenges for network administrators to maintain visibility and ensure reliable network performance. Consequently, it is becoming increasingly critical to continuously monitor network traffic to understand traffic flows, preserve performance, and ensure the reliability and security of modern networks.

One critical component of traffic monitoring is feature extraction, which captures key characteristics of network flows and enables accurate traffic classification, anomaly detection, and behavioral analysis. Flow-level feature extraction, in particular, has become the dominant approach due to its ability to generate informative summaries of network patterns [3]. It allows network administrators to efficiently characterize large volumes of traffic and understand holistic traffic behavior without deeply inspecting individual packets. Several tools have been proposed for flow feature extraction, such as CICFlowMeter, ICSFlowGenerator, LycoSTand and LiteEx [3]–[6].

Accurately identifying network flows is essential for reliable flow feature extraction. Existing flow feature extraction tools mainly rely on 5-tuple attributes (i.e., source IP, destination IP, source port, destination port, and protocol), where packets sharing the same 5-tuple are grouped into a single flow. While the mechanism is effective to aggregate packets exchanged between two endpoints, it may lead to fragmented or incomplete flow reconstruction. Specifically, evasive techniques, such as IP spoofing, port rotation, or multi-session tunneling can intentionally distribute packets of a single activity across multiple apparent flows to avoid being correctly aggregated by the 5-tuple method. In addition, conventional flow feature extraction is typically performed after a flow has ended or timed out. The reliance on post-completion analysis limits real-time network feature extraction and mitigation of ongoing attacks.

To address these limitations, we propose Hi-Fi Flow, a novel real-time high-granularity flow identification and feature extraction tool. Instead of relying solely on location-based 5-tuple attributes, Hi-Fi Flow incorporates spatial attributes (e.g., endpoint information), structural attributes (e.g., packet ordering, sequence patterns), and temporal attributes (e.g., inter-arrival times, temporal proximity) to capture correlations among packets that belong to the same underlying activity. By jointly modeling spatial, structural, and temporal dependencies, Hi-Fi Flow can accurately identify network flows even under dynamic and evasive behaviors. Additionally, Hi-Fi Flow incrementally updates flow-level features as packets arrive to enable real-time feature extraction and facilitate timely detection and mitigation of ongoing network attacks.

To achieve these goals, Hi-Fi Flow addresses the following essential technical challenges.

Real-time network flow identification: Traditional flow identification aggregates packets into flows based solely on location-derived attributes, which may overlook sophisticated or evasive packet behaviors. Hi-Fi Flow overcomes this limitation by introducing a dual-encoder architecture that includes a spatial encoder and a behavior encoder. The spatial encoder

2158-3927/26/$31.00 ©2026 IEEE
DOI 10.1109/DSN69566.2026.00052

captures spatial dependencies by encoding endpoint-related packet attributes, while the behavior encoder learns structural and temporal consistency patterns across packet sequences. Both encoders employ a self-attention mechanism to model long-range dependencies and capture correlations that conventional 5-tuple based flow grouping fails to recognize. In addition, Hi-Fi Flow supports real-time incremental flow identification. It allows the system to dynamically aggregate packets as they arrive and continuously update discriminative flow-level features, facilitating timely detection of ongoing attacks rather than relying on post-flow completion analysis.

Data-efficient flow identification training: Flow identification training typically requires large, manually labeled datasets, which are costly and labor-intensive to create. To reduce dependence on extensive labeling, Hi-Fi Flow adopts a few-shot triplet learning strategy to train the flow identification model with only a small labeled dataset. Instead of relying on absolute class labels, the model learns relative similarity relationships among packet-flow samples by pulling embeddings from the same flow closer together and pushing embeddings from different flows farther apart. This training strategy improves flow discrimination capability while significantly reducing annotation overhead, enabling accurate flow identification with limited labeled data.

In summary, the work makes the following key contributions to the field of network flow analysis:

1) We develop Hi-Fi Flow, a high-granularity flow representation system that goes beyond static 5-tuple flow construction by modeling fine-grained spatial, temporal, and structural correlations among packets.
2) We design a dual self-attention encoder that combines spatial and behavioral dependencies to enable robust, real-time packet-flow association under dynamic routing, evasion, and adversarial conditions.
3) We introduce a few-shot triplet learning strategy that enables accurate flow identification with very limited labeled data and substantially reduces manual annotation overhead.
4) Extensive evaluations show that Hi-Fi Flow consistently delivers near-perfect flow identification with few labeled samples, and its features enable state-of-the-art intrusion detection across diverse datasets, outperforming all baseline flowmeters.
5) We will publicly release Hi-Fi Flow and supporting datasets to enable further research in high-resolution flow analysis and intrusion detection.

II. Related Work

Flow feature extraction has become a core component of modern network monitoring and intrusion detection. It aggregates raw packets into semantically meaningful flow records for efficient traffic analysis and security detection [7].

There exist multiple tools developed for flow-level feature extraction. On the research side, tools such as CICFlowMeter, ICSFlowGenerator, LycoSTand, and LiteEx provide enriched flow representations specifically designed to support machine-learning based traffic classification and anomaly detection [3]–[6]. On the industrial side, standardized flow exporters including NetFlow, IPFIX, and sFlow as well as advanced software monitoring frameworks such as Zeek, Tstat, YAF, NFDUMP, and pmacct are widely used for real-world traffic monitoring and feature extraction [8]–[14].

Nevertheless, existing flow feature extraction tools share two fundamental limitations. First, they solely rely on 5-tuple rules to group packets into flows. Such a heuristic grouping mechanism often fails to correctly aggregate packets generated by dynamic or evasive behaviors (e.g., IP spoofing, port rotation, multi-session tunneling), resulting in fragmented or incomplete flow records [15]–[17]. Second, many tools extract flow features only after a flow has ended via explicit session termination (e.g., protocol flag) or timeout. The post-completion processing prevents real-time flow construction and timely detection of ongoing malicious activity.

To this end, we propose Hi-Fi Flow, a fine-grained, learning-based flow feature extraction tool that correlates packets in real time to achieve more accurate and robust flow construction. Unlike traditional 5-tuple based approaches, Hi-Fi Flow jointly models spatial, structural, and temporal relationships to identify packets belonging to the same underlying activity.

III. Threat Model

We consider a heterogeneous network composed of diverse devices (e.g., servers, user terminals, and IoT nodes) that continuously generate large volumes of network traffic across multiple protocols. We assume adversaries capable of generating diverse malicious traffic targeting network hosts and services, ranging from large-scale botnet driven attacks to stealthy, distributed activities that use IP rotation, spoofing, or multi-session tunneling to appear low-volume or intermittent. Hi-Fi Flow is deployed at trusted observation points (e.g., Internet gateways, border routers), where it passively monitors packet and flow metadata without disrupting normal operations. Its objective is to accurately aggregate packets into activity-based flows and extract robust, discriminative flow-level features that can be used for effective attack detection.

IV. System Overview

In this section, we provide a brief overview of the overall design of the proposed Hi-Fi Flow system.

Hi-Fi Flow is designed to perform real-time, high-granularity flow identification and feature extraction from raw network traffic. As illustrated in Figure 1, the system adopts an incremental processing architecture that dynamically associates incoming packets with active flows and continuously updates flow-level features. In particular, the proposed tool includes four components:

User-Friendly Graphical User Interface (GUI) The tool features an intuitive GUI that serves as the primary control panel, allowing users to activate real-time traffic processing, configure system parameters, and manage flow identification and feature extraction tasks. The dashboard provides live

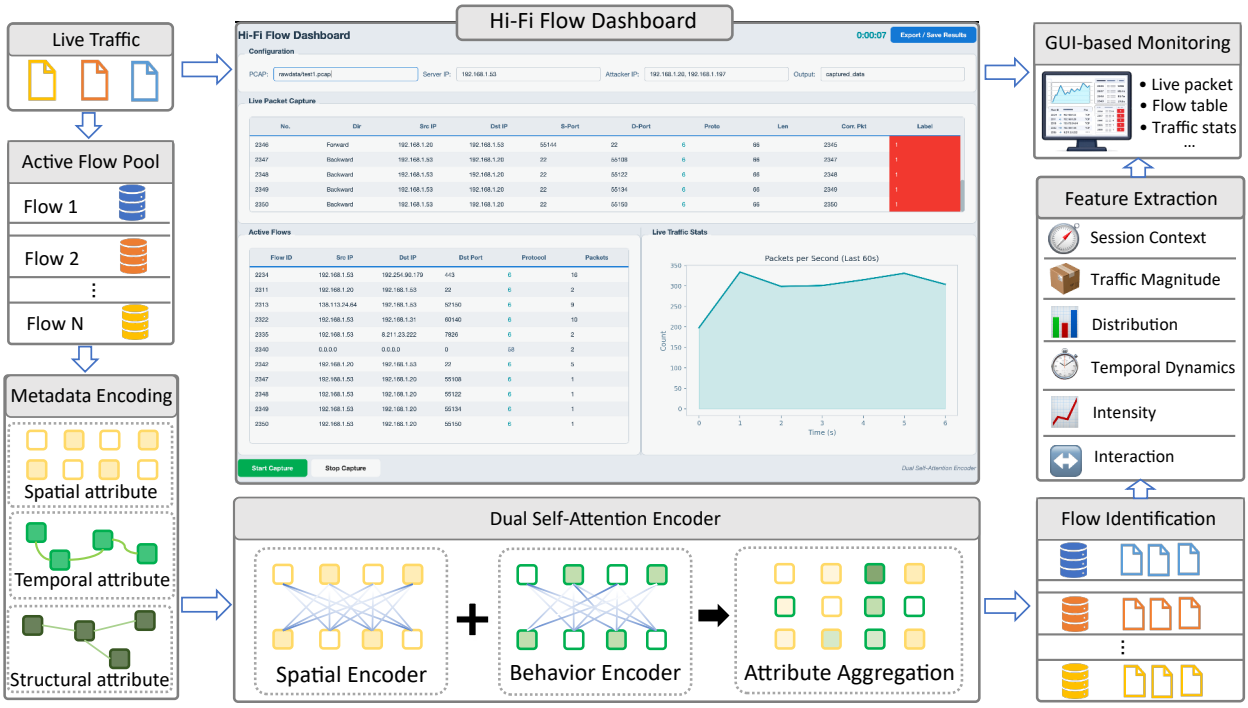


Fig. 1. System Overview of Hi-Fi Flow

statistics such as packet counts, traffic composition, and aggregated behavioral metrics, giving administrators immediate visibility into network conditions as shown in Figure 1.

High-Granularity Flow Identification This module enables accurate real-time flow identification by jointly modeling spatial, structural, and temporal correlations among packets. *Traffic metadata embedding* first converts raw packet attributes into structured vectors, which are then processed by a dual-encoder architecture: a *spatial encoder* that captures endpoint-level relationships, and a *behavior encoder* that learns temporal and structural dependencies across packet sequences. By combining these complementary representations, Hi-Fi Flow reliably associates packets with their true flows even under dynamic routing changes or evasive behaviors.

Learning-Efficient Flow Identification Training Hi-Fi Flow minimizes manual labeling requirements through a few-shot triplet learning strategy. The triplet-network architecture enables the model to learn discriminative packet-flow similarity patterns from a very limited labeled dataset. By comparing anchor, positive, and negative packet-flow samples, the model learns an embedding space in which packets from the same flow are closer together while packets from different flows are separated. This design reduces annotation overhead while maintaining accurate flow identification under limited supervision.

Real-Time Flow Feature Extraction This module incrementally computes fine-grained flow features as packets arrive. Once incoming packets are associated with their corresponding flows, Hi-Fi Flow continuously updates statistical and behavioral features for each active flow without waiting for flow termination or timeout. The extracted features cover session context, traffic magnitude, distributional characteristics, temporal dynamics, intensity, and interaction state, providing rich representations for traffic analysis and downstream intrusion detection.

V. High-Granularity Flow Identification

In this section, we present the detailed design of the high-granularity flow identification module in Hi-Fi Flow. The goal of the module is to capture fine-grained spatial, temporal, and structural correlations among packets in order to construct highly discriminative flow representations for feature extraction.

Hi-Fi Flow maintains an **active flow pool** that stores the real-time state of all ongoing flows. When a new packet arrives, the tool compares it against all active flows to determine whether the packet should be associated with an existing flow or used to create a new one. To facilitate accurate comparison, **traffic metadata embedding** converts raw packet attributes into structured spatial, temporal, and structural representations. These embeddings are fed into the **dual encoder** for flow association, which uses a dual self-attention mechanism to compute fine-grained correlations between incoming packets and candidate flows.

A. Active Flow Pool

The proposed tool maintains an active flow pool to track ongoing flows in real time. Instead of buffering raw packets,

each flow entry records a lightweight summary of spatial coherence (e.g., IP deviation, port and protocol alignment), temporal regularity (e.g., inter-arrival time statistics and timing deviations), and structural correlation (e.g., standardized deviations of payload, packet, and header lengths) to improve processing efficiency and reduce memory overhead. To further improve scalability, each flow is assigned an adaptive timeout value τ_f based on its historical inter-packet timing behavior. A flow is considered inactive and removed when no new packets are associated with it for a period longer than τ_f.

B. Metadata Embedding

To enable accurate packet-flow association, raw packet attributes are transformed into structured metadata representations capturing spatial, temporal, and structural relationships. These representations allow packet-flow correlations to be computed efficiently without exhaustive comparison of raw attributes.

1) Spatial Embedding: Spatial consistency measures how closely the endpoint identifiers of an incoming packet (e.g., source and destination IP addresses, ports, and protocol) align with the established communication pattern of an existing flow, which summarizes the observed IP address patterns, port distributions, protocol type, and deviation statistics derived from previously associated packets. The spatial embedding is defined as:

$$\mathbf{S} = [\Delta_{IP}, \Delta_{Pt}, \Delta_{Pr}],$$

where Δ_{IP} is the IP deviation computed by segment-wise XOR between packet and flow address octets, Δ_{Pt} is the port similarity represented as a binary match indicator, Δ_{Pr} is the protocol similarity encoded as a single-bit deviation.

2) Temporal Embedding: Temporal consistency characterizes whether a packet arrival time aligns with the established temporal pattern of a flow, which captures observed inter-arrival intervals, temporal variance, and deviation statistics derived from previously associated packets. The temporal embedding is computed as:

$$\mathbf{T} = [\overline{\Delta_t},\, \sigma_t,\, Z_t],$$

where the $\overline{\Delta_t}$ describes typical packet spacing for the flow, σ_t tracks temporal variability (e.g., bursty vs. periodic behavior), and Z_t is the z-score based deviation which quantifies how unusual the current packet timing is compared to past timing patterns [18].

3) Structural Embedding: Structural consistency evaluates whether the framing characteristics of an incoming packet (payload size, packet length, and header length) match the statistical profile of an existing flow. The structural embedding is computed as

$$\mathbf{H} = [Z_{Pl}, Z_{Pk}, Z_{Hd}],$$

where each entry is a standardized structural deviation of payload size, frame length, and header size.

C. Dual-Encoder for Flow Identification

To accurately determine whether an incoming packet belongs to a candidate flow, Hi-Fi Flow employs a dual-encoder architecture, one modeling spatial relationships and the other modeling temporal-structural behaviors.

1) Spatial Encoder: For each incoming packet, the spatial encoder processes spatial embedding $\mathbf{S}$ and transforms it into Query, Key, and Value matrices:

$$\mathbf{Q}_s = \mathbf{S}\mathbf{W}_s^Q, \qquad \mathbf{K}_s = \mathbf{S}\mathbf{W}_s^K, \qquad \mathbf{V}_s = \mathbf{S}\mathbf{W}_s^V,$$

where $\mathbf{W}_s^Q$, $\mathbf{W}_s^K$, and $\mathbf{W}_s^V$ are learned projection matrices. The spatial correlation embedding is computed using scaled dot-product attention:

$$\mathbf{A}_s = \text{Softmax}\left(\frac{\mathbf{Q}_s\mathbf{K}_s^\top}{\sqrt{d_s}}\right)\mathbf{V}_s,$$

where d_s is the length of the projected vectors.

2) Behavior Encoder: The behavior encoder jointly processes temporal and structural metadata $\mathbf{B} = [\mathbf{T}|\mathbf{H}]$ and similarly generates:

$$\mathbf{Q}_b = \mathbf{B}\mathbf{W}_b^Q, \qquad \mathbf{K}_b = \mathbf{B}\mathbf{W}_b^K, \qquad \mathbf{V}_b = \mathbf{B}\mathbf{W}_b^V.$$

Behavior correlation embedding is then computed as:

$$\mathbf{A}_b = \text{Softmax}\left(\frac{\mathbf{Q}_b\mathbf{K}_b^\top}{\sqrt{d_b}}\right)\mathbf{V}_b.$$

The outputs $\mathbf{A}_s$ and $\mathbf{A}_b$ are concatenated to form a packet-flow embedding. We then adopt a multilayer perceptron (MLP) to compute the packet-flow similarity score for each active flow in the pool. The incoming packet is assigned to the flow with the highest similarity score if the score exceeds a predefined decision threshold. Otherwise, the system initializes a new flow entry in the pool.

VI. Learning-Efficient Flow Identification Training

Hi-Fi Flow reduces manual labeling overhead using the **few-shot triplet learning** strategy. Triplet learning enables the model to learn discriminative packet-flow similarity from a limited labeled dataset by comparing relative relationships among packet-flow samples rather than relying on large-scale labeled training data.

A. Few-Shot Triplet Learning

Hi-Fi Flow adopts a few-shot triplet learning strategy to train the flow identification model with minimal labeled data. Instead of predicting absolute class labels, the model learns relative similarity by comparing triplets: an anchor packet-flow sample, a positive sample from the same flow, and a negative sample from a different flow. The goal is to make embeddings of similar flows closer together and embeddings of dissimilar flows farther apart. The triplet loss is defined as:

$$L_{\text{triplet}} = \lambda_{ap}\, d_{ap} + \lambda_{an}\, \max\left(0, m_{an} - d_{an}\right),$$

where d_{ap} and d_{an} represent embedding distances between anchor-positive and anchor-negative pairs, respectively. We

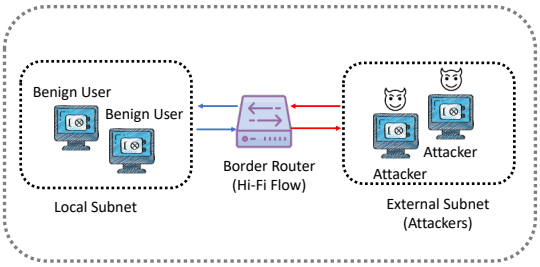


Fig. 2. Experiment Testbed

set the margin m_{an} to 1 to ensure sufficient separation between negative samples, while the weighting coefficients $\lambda_{ap} = \lambda_{an} = 0.5$ balance the contribution of each term. The strategy produces a compact and discriminative embedding space, allowing accurate flow identification with only a few labeled examples.

VII. Flow Feature Extraction

Once packets are accurately associated with their corresponding flows, Hi-Fi Flow performs fine-grained feature extraction to generate rich behavioral and statistical representations.

Hi-Fi Flow computes 110 flow-level features, grouped into six major categories:

- **Session Context Features** describe the overall communication session, such as flow start time, end time, duration, and traffic direction.
- **Traffic Magnitude Features** quantify how much data is exchanged within a flow, including total bytes, total packets, and directional traffic counts.
- **Distributional Features** characterize how packet and payload sizes are distributed within the flow, including minimum, maximum, mean, and variance of packet lengths, header sizes, and payload sizes in both directions.
- **Temporal Dynamics Features** describe the timing pattern of packet transmission, such as inter-arrival times, jitter, burstiness, burst duration, and idle periods.
- **Intensity Features** measure how aggressively traffic is delivered over time, including packets per second, bytes per second, and throughput fluctuations.
- **Interaction State Features** summarize transport-layer signaling and connection behavior, such as TCP flag counts including SYN, ACK, FIN, and RST.

VIII. Experimental Evaluation

A. Experiment Setup

The experiment is conducted on a controlled institutional network managed by a border router implemented on a network hypervisor as shown in Figure 2. The router segregates the network into two subnets: a local subnet with roughly 50 active IP addresses (e.g., user workstations and departmental servers) and an external subnet that hosts controlled attacker machines. The router manages all inter-subnet routing and enforces access control to isolate the traffic. We deploy Hi-Fi Flow on the border router to monitor inbound and outbound traffic of the local network. The setup enables passive, full-fidelity traffic capture over a 1 Gbps link and allows real-time flow identification and feature extraction under realistic, mixed benign and adversarial conditions.

The local subnet emulates realistic network activities, maintaining $10 \sim 50$ concurrent TCP/UDP sessions that include interactive traffic (e.g., web, SSH) and background research workloads. Meanwhile, attacker hosts generate 10 scripted attack variants. Table I lists representative attacks used as a realistic benchmark to evaluate the proposed flow feature extraction tool. Beyond standard attack behaviors, we incorporate evasive techniques that deliberately distribute malicious traffic across multiple sessions, causing traditional 5-tuple based aggregation to fragment a single logical activity into separate flows. For example, in the *Slowloris attack*, we split long-lived HTTP connections across multiple ephemeral source ports, creating several parallel low-rate sessions. In the *Port Scan attack*, we apply rapid port rotation to hide scanning patterns and force conventional flowmeters to treat related probes as independent flows. These adaptations create a more challenging and realistic evaluation environment for high-granularity flow reconstruction.

To evaluate Hi-Fi Flow beyond the scale of our institutional testbed, we further validate the proposed tool using three large-scale public benchmark datasets, **UNSW-NB15**, **CIC-IDS2017**, and **ICS Pcap**. These datasets comprise over **150 GB of raw network traffic** and millions of flow records, spanning diverse attack scenarios and complex enterprise-scale network environments. The extended evaluation allows us to assess the scalability and robustness of Hi-Fi Flow under more diverse and data-intensive traffic conditions.

TABLE I
Attack Variants and Evasive Adaptations

#	Attack	Evasive Adaptation and 5-tuple Manipulation
1	Slowloris	**Multi-session**: Traffic split across 5–10 ephemeral ports to avoid per-flow rate limits.
2	Port Scan	**Port Rotation**: Rapid switching of src/dst ports to mask the vertical/horizontal scan.
3	SYN Flood	**IP Spoofing**: Forged source IP octets via XOR logic to simulate a distributed attack.
4	UDP Flood	**IP Spoofing**: Randomized source addresses to shatter flow-based tracking and statistics.
5	SQL Inj.	**Fragmentation**: Spreading malicious query strings across multiple sub-flows to hide signatures.
6	XSS Inj.	**Encapsulation**: Hiding scripts within multi-session tunnels to bypass static 5-tuple inspection.
7	FTP Brute	**Distributed Sessions**: Login attempts split across varied TCP flows to evade threshold alerts.
8	SSH Brute	**Distributed Sessions**: Mixing malicious packets with benign background traffic across ports.
9	NTP Amp.	**IP Spoofing**: Forging victim's IP as source to redirect high-volume amplified responses.
10	SNMP Amp.	**IP Spoofing**: Forging source IPs to weaponize SNMP servers for reflection and amplification.

TABLE II
EFFECTIVENESS OF FLOW IDENTIFICATION: STABILITY ANALYSIS ACROSS SAMPLE SIZES (20 TRIALS) IN(%)

Samples	Stat.	Benign	Slowloris	Port Scan	SYN Flood	UDP Flood	SQL Inj.	XSS Inj.	FTP Brute	SSH Brute	NTP Amp.	SNMP Amp.	Overall
	Max	65.42	62.83	48.94	54.56	58.31	46.92	45.31	39.07	44.81	50.14	40.76	**50.64**
10	Med	58.15	54.12	41.25	46.88	50.42	38.45	37.12	31.54	36.92	42.15	32.41	**42.67**
	Min	52.33	47.56	33.12	39.45	43.18	31.05	29.87	24.12	28.56	34.22	25.13	**35.33**
	Max	92.84	91.58	79.99	87.85	89.98	74.92	73.31	74.07	73.80	79.12	68.48	**80.54**
20	Med	89.12	87.45	75.12	83.21	85.34	70.12	68.45	69.21	68.95	74.32	63.56	**75.89**
	Min	86.45	83.12	71.45	79.56	81.22	66.85	64.12	65.34	64.78	70.15	59.42	**72.04**
	Max	94.88	93.57	91.97	93.85	94.98	90.34	92.32	91.07	92.31	93.15	90.78	**92.65**
30	Med	93.12	91.22	89.45	91.56	92.84	88.12	90.15	89.12	90.34	91.24	88.56	**90.52**
	Min	91.45	89.15	87.12	89.42	90.56	86.45	88.23	87.45	88.12	89.56	86.21	**88.52**
	Max	95.91	94.27	94.97	95.99	96.81	94.05	95.41	96.89	95.93	96.82	95.97	**95.73**
40	Med	95.22	93.51	94.12	95.22	96.15	93.21	94.72	96.21	95.12	96.14	95.23	**94.98**
	Min	94.76	92.88	93.45	94.61	95.54	92.51	94.12	95.66	94.45	95.52	94.62	**94.38**
	Max	96.45	95.28	97.97	98.99	99.98	98.31	98.96	97.89	98.24	99.82	97.97	**98.17**
50	Med	96.28	95.11	97.79	98.80	99.71	98.00	98.78	97.71	98.03	99.71	97.79	**97.97**
	Min	96.12	94.97	97.66	98.62	99.40	97.71	98.58	97.56	97.87	99.56	97.66	**97.73**
Full		99.12	98.51	99.99	99.99	99.99	99.75	98.95	98.89	99.34	99.23	99.31	**99.37**

B. Evaluation of Flow Identification Effectiveness

We evaluate the effectiveness of the flow identification model under limited supervision. Our goal is to demonstrate that Hi-Fi Flow can learn accurate packet-flow association using only a small number of labeled flow instances per attack.

1) Few-Shot Training Setup: We conduct a few-shot learning experiment using 10, 20, 30, 40, and 50 labeled flow instances per attack type. The model is trained solely with these few-shot instances using the proposed triplet learning architecture. This setup evaluates whether Hi-Fi Flow can learn discriminative packet-flow similarity patterns with limited labeled data.

2) Effectiveness of the Dual-Encoder Architecture: We first evaluate the effectiveness of the dual-encoder architecture. We compare three configurations for flow identification: (1) using only the Spatial Encoder, (2) using only the Behavior Encoder, and (3) using the full Dual Attention Encoder. Each model is trained under the same few-shot setup, using limited labeled samples per attack, and evaluated on identical test traffic.

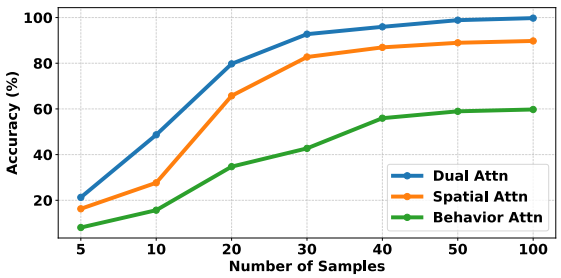


Fig. 3. Effectiveness of the Dual-Encoder

Figure 3 shows a consistent accuracy gap among the three configurations. The dual encoder achieves the highest accuracy across all sample sizes and reaches near-perfect identification accuracy, close to 100%, with only 50 labeled samples per attack. In contrast, the spatial encoder alone reaches a maximum accuracy of roughly 90%, while the behavior encoder alone stabilizes at around 60%. This performance gap demonstrates that spatial similarity or temporal-structural consistency alone is insufficient to fully characterize complex and evasive flows.

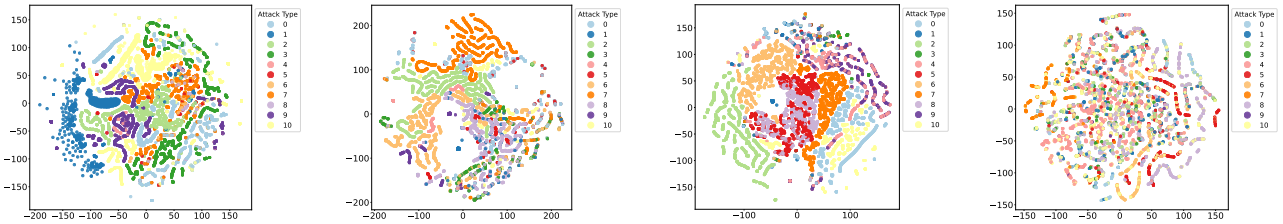
By jointly learning endpoint-level spatial dependencies and temporal-structural patterns, the dual encoder provides substantially more reliable packet-flow association and stronger discrimination among different attack behaviors. The results validate the necessity of integrating both types of information into the flow identification process, especially in dynamic or adversarial network environments.

3) Few-Shot Identification Accuracy Evaluation: Table II summarizes flow identification accuracy, defined as the percentage of packets correctly associated with their underlying activity-level flows, across 10 attack scenarios under varying few-shot training sizes. For each sample size, we conduct 20 independent trials using randomly selected labeled training packets and record the maximum, median, and minimum accuracy to evaluate robustness and stability. The results show that Hi-Fi Flow effectively reconstructs activity-level flows under both high-volume attacks and stealthy, distributed behaviors. In the extremely low-shot setting with 10 samples, accuracy varies noticeably due to limited supervision. However, performance improves rapidly as more labeled samples are used. When the number of labeled samples increases to 30, the median flow identification accuracy exceeds 90% for most scenarios, and the variance across trials narrows substantially. With 50 samples, Hi-Fi Flow achieves near-optimal and highly stable performance, with median accuracy typically above 97%-99% and very small max-min gaps. Overall, the results demonstrate that Hi-Fi Flow can reliably reconstruct coherent activity-level flows with limited labeled data across diverse attack scenarios.

4) Few-Shot Identification Accuracy Across Different Datasets: To further evaluate the generalizability and scalability of Hi-Fi Flow, we conduct extensive experiments across four datasets of varying scale and complexity. As shown in Table III, flow identification performance improves consistently as the number of labeled samples increases. In the extremely low-shot setting (10 samples), F1 scores range from 35% to 51%. With 50 labeled samples, both F1 and accuracy

TABLE III
EFFECTIVENESS OF FLOW IDENTIFICATION ACROSS DIFFERENT RAW DATASETS IN(%)

Samples	CICIDS2017 Dataset				Hi-Fi Flow Dataset				ICS Pcap Dataset				UNSW-NB15 Dataset			
	F1	ACC	PRE	REC	F1	ACC	PRE	REC	F1	ACC	PRE	REC	F1	ACC	PRE	REC
10	41.96	47.08	40.03	44.07	50.57	49.17	50.53	50.60	43.06	47.32	42.23	43.97	35.55	45.46	33.92	37.33
20	85.43	81.88	85.22	85.63	81.14	79.31	81.54	80.75	79.33	80.13	79.12	79.73	77.52	80.94	76.69	78.71
30	88.01	86.05	87.73	88.29	93.32	92.50	93.64	93.00	86.31	85.68	86.08	86.54	88.79	90.17	87.62	90.00
40	92.96	92.21	92.83	93.09	95.78	95.71	95.87	95.69	92.71	93.46	92.11	93.32	91.67	94.39	90.20	93.19
50	97.94	98.39	97.97	97.91	98.24	98.34	98.10	98.38	97.43	97.54	97.35	97.51	97.34	97.68	97.28	97.41
Full	99.98	99.97	99.98	99.98	99.78	99.40	99.77	99.79	99.37	99.94	99.37	99.37	99.44	99.10	99.43	99.46



(a) Hi-Fi Flow (b) CIC Flowmeter (c) LycoSTand Flowmeter (d) ICS FlowGenerator

0: Slowloris; **1**: Port Scan; **2**: SYN Flood; **3**: UDP Flood; **4**: NTP Amplification; **5**: XSS Injection; **6**: FTP Brute Force; **7**: SSH Brute Force; **8**: SQL Injection; **9**: SNMP Amplification; **10**: Benign

Fig. 4. T-SNE Comparison of Four Network Extraction Tools

exceed 97% across all datasets, demonstrating robust few-shot flow reconstruction even under large-scale, data-intensive conditions. These results indicate that the few-shot learning strategy enables the model to quickly learn discriminative flow features with minimal supervision, allowing accurate activity-level flow reconstruction and laying a solid foundation for attack detection.

C. Visualization of Flow Feature Separability

We compare the flow feature representations generated by Hi-Fi Flow with those extracted by conventional 5-tuple based flow feature tools, including CICFlowMeter, LycoSTand, and ICSFlowGenerator. To assess separability and discriminative capability, we apply t-Distributed Stochastic Neighbor Embedding (t-SNE) to project the high-dimensional flow embeddings into a two-dimensional space for visualization and qualitative analysis.

Figure 4 exhibits substantial differences in feature quality across the four tools. As shown in Figure 4(a), Hi-Fi Flow produces tight, well-defined, and clearly separated clusters for all attack classes. Different attack behaviors occupy distinct and non-overlapping regions in the 2D space, indicating that Hi-Fi Flow captures fine-grained behavioral characteristics and produces highly discriminative flow embeddings for accurate attack detection.

Figures 4(b) and 4(c) show that CICFlowMeter and LycoSTand also achieve good clustering for most attack types. Nevertheless, several classes, such as Attack 3 and 8 for CICFlowMeter, Attack 5 and 8 for LycoSTand form mixed clusters with substantial overlap, making them difficult to distinguish. In addition, the termination behavior of 5-tuple based flow extractors causes variation in the number of flows, which further affects feature representation quality. For instance, Attack 0 and 1 in CICFlowMeter produce only a small number of valid flow records, resulting in limited statistical and behavioral features that may be insufficient to comprehensively characterize the attacks.

In contrast, Figure 4(d) reveals the weakest separability for ICSFlowGenerator. Feature points across all attack types are heavily mixed, forming no clear clusters or decision boundaries. This indicates that ICSFlowGenerator extracted features lack the discriminatory detail required to separate diverse attack behaviors.

D. Comparative Evaluation

This section evaluates the effectiveness of Hi-Fi Flow for attack identification and malicious traffic detection. We first examine how well the extracted flow features generalize across different datasets. We then compare Hi-Fi Flow with other flow feature extraction tools to quantify the benefits of high-granularity flow identification and feature extraction.

1) Detection Accuracy Across Different Raw Datasets: Table IV evaluates the robustness of the Hi-Fi Flow by applying the same feature extraction process to four distinct network datasets: CICIDS2017, Hi-Fi Flow Data, ICS Pcap, and UNSW-NB15 [19]–[21]. Across all datasets, most classical machine-learning models (DecisionTree, RandomForest, AdaBoost, XGBoost, and LightGBM) achieve consistently good detection performance, with AUROC, AUPRC, and Accuracy frequently exceeding 99% on both CICIDS2017 and Hi-Fi Flow Data. Even on more heterogeneous traffic such as UNSW-NB15, the Hi-Fi features maintain strong performance;

TABLE IV
DETECTION ACCURACY IN(%) OF HI-FI FLOW ACROSS DIFFERENT RAW DATASETS

Model	CICIDS2017 Dataset						Hi-Fi Flow Dataset						ICS Pcap Dataset						UNSW-NB15 Dataset					
	AUROC	AUPRC	F1	ACC	PRE	REC	AUROC	AUPRC	F1	ACC	PRE	REC	AUROC	AUPRC	F1	ACC	PRE	REC	AUROC	AUPRC	F1	ACC	PRE	REC
DecisionTree	99.68	98.84	99.70	99.82	99.84	99.57	99.37	97.07	98.69	98.58	99.15	98.24	99.84	99.69	99.36	99.31	99.54	99.18	99.69	98.39	99.07	99.76	99.77	98.38
RandomForest	98.84	98.12	99.48	99.78	99.79	99.17	99.81	99.56	97.94	98.01	98.16	97.72	99.76	99.58	99.45	99.31	99.78	99.12	99.68	97.49	98.34	99.54	99.31	97.39
AdaBoost	98.74	98.19	99.34	98.70	99.71	98.97	99.91	98.67	99.26	98.07	99.90	98.63	99.91	99.71	99.60	99.98	99.12	99.08	99.99	99.90	98.33	99.14	99.19	97.49
XGBoost	98.95	98.34	99.57	98.99	99.68	99.47	99.90	99.72	99.04	99.18	99.30	98.78	99.88	99.61	99.41	97.19	99.45	99.37	99.18	99.04	99.48	99.85	99.59	99.38
LightGBM	98.99	98.99	99.39	97.80	99.84	98.95	99.50	99.03	99.23	99.53	99.65	98.82	97.99	96.88	97.93	99.99	98.19	97.68	99.19	98.84	99.40	99.90	99.49	99.31
CatBoost	99.98	98.99	99.42	98.85	99.87	98.97	99.88	99.58	98.46	99.09	99.15	97.78	99.99	99.91	98.10	96.99	98.19	98.02	99.34	97.88	97.63	98.99	98.12	97.15
MLP	98.95	97.99	99.32	97.68	99.73	98.92	99.56	99.24	99.04	98.29	99.44	98.64	99.98	99.63	97.30	98.89	97.68	96.92	99.70	99.27	98.45	97.95	98.76	98.15
LinearSVC	98.73	97.92	97.25	97.45	97.78	96.73	98.75	97.92	97.83	98.98	98.76	96.90	97.99	97.68	98.20	98.04	98.99	97.43	99.37	98.16	96.02	99.72	97.36	94.74
SVCRBF	98.76	96.88	98.66	98.38	99.68	97.65	98.34	97.50	99.41	96.86	99.92	98.92	98.68	97.94	98.25	97.53	98.56	97.95	99.36	97.99	95.38	96.76	96.59	94.19

TABLE V
DETECTION ACCURACY IN(%) ACROSS DIFFERENT FLOW FEATURE EXTRACTION TOOLS UNDER SAME RAW DATASET

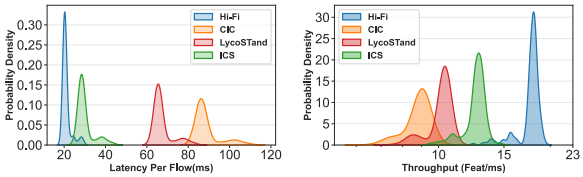
Model	Hi-Fi Flow						CICFlowMeter						LycoSTand						ICSFlowGenerator					
	AUROC	AUPRC	F1	ACC	PRE	REC	AUROC	AUPRC	F1	ACC	PRE	REC	AUROC	AUPRC	F1	ACC	PRE	REC	AUROC	AUPRC	F1	ACC	PRE	REC
DecisionTree	98.95	98.19	97.05	99.99	97.12	96.99	96.32	95.94	94.98	95.90	95.01	94.94	91.42	90.96	89.94	91.13	90.26	89.63	91.63	91.24	90.38	90.71	90.62	90.14
RandomForest	98.41	98.24	98.19	99.98	98.29	98.09	95.81	95.36	93.64	94.11	94.16	93.12	90.99	90.08	90.07	90.19	90.19	89.95	90.38	90.19	90.81	91.83	91.19	90.44
AdaBoost	99.42	98.54	98.21	99.90	98.46	97.97	95.81	95.57	94.05	94.30	94.90	93.23	91.81	91.72	91.21	91.84	91.91	90.53	92.80	91.30	91.21	90.29	91.55	90.87
XGBoost	99.31	97.34	98.07	98.19	98.19	97.95	94.10	92.29	91.97	95.12	92.24	91.71	92.48	92.31	90.16	90.21	90.19	90.14	91.19	90.34	90.61	91.38	90.91	90.32
LightGBM	98.99	98.88	98.59	98.91	98.99	98.19	95.43	94.93	94.23	94.53	94.65	93.82	92.35	91.18	91.58	90.96	91.93	91.23	90.90	90.83	90.27	91.88	90.32	90.22
CatBoost	97.98	97.69	97.48	96.96	97.98	96.98	96.49	95.28	94.27	93.94	95.91	92.63	92.12	91.81	90.35	90.79	90.39	90.31	91.21	90.38	90.58	91.42	91.04	90.12
MLP	98.95	98.04	99.92	97.99	99.97	99.87	95.56	94.91	94.46	94.29	94.13	93.62	93.58	92.63	91.54	91.87	91.67	91.42	91.29	90.92	87.86	90.30	88.21	87.51
LinearSVC	98.83	97.81	99.26	98.95	99.68	98.85	95.75	93.92	93.73	95.98	93.76	93.70	91.48	91.38	91.82	90.13	92.32	91.32	90.02	89.66	86.77	90.90	87.64	85.92
SVCRBF	98.96	97.98	98.28	98.79	98.83	97.73	89.38	88.41	92.82	95.51	92.92	92.72	92.88	92.34	91.25	88.94	91.56	90.95	88.65	87.11	89.78	88.01	90.08	89.48

for instance, XGBoost achieves 99.85% accuracy, 99.48% F1, and 99.59% precision. The results indicate that Hi-Fi Flow produces highly expressive feature representations that remain accurate and resilient across heterogeneous network environments.

2) Detection Accuracy Across Different Flow Feature Extraction Tools: Table V compares the attack detection performance when the same raw network dataset is processed using four different flow feature extraction tools: Hi-Fi Flow, CICFlowMeter, LycoSTand, and ICSFlowGenerator. As shown, machine learning models trained on Hi-Fi Flow features generally achieve the best detection results, with several models approaching high AUROC, AUPRC, F1, Precision, and Recall scores. Traditional flowmeters such as CICFlowMeter and LycoSTand still deliver competitive performance, however, slightly lower accuracy and feature separability are observed, particularly for complex or mixed traffic conditions. For example, RandomForest achieves 99.98% accuracy and 98.19% F1 using Hi-Fi Flow features, but its accuracy decreases to 94.11% (F1 = 93.64%) with CICFlowMeter features, and further to 90.19% (F1 = 90.07%) when using LycoSTand features. ICSFlowGenerator yields the weakest performance overall, with noticeable metric reductions and reduced class distinction across models. Overall, the comparison suggests that Hi-Fi Flow tends to produce more informative flow features, which in turn enhance downstream intrusion detection compared to several widely used legacy extraction tools.

E. Performance Overhead

We further evaluate the computational overhead of Hi-Fi Flow to assess its practicality for real-time deployment. We compare the flow extraction latency and feature processing throughput with several representative baseline tools. As



(a) Latency (b) Throughput

Fig. 5. Performance Overhead Evaluation

shown in Figure 5(a), Hi-Fi Flow achieves consistently lower latency because it performs incremental flow construction and updates flow statistics as packets arrive instead of waiting for flow completion or timeout events before exporting features. Figure 5(b) shows that Hi-Fi Flow also maintains higher throughput under sustained traffic load, as it computes features on compact flow summaries rather than buffering and reprocessing full packet traces. Overall, Hi-Fi Flow delivers efficient, low-latency, and high-throughput feature extraction suitable for real-world high-speed network monitoring.

IX. CONCLUSION

Hi-Fi Flow provides a real-time, high-granularity flow identification and feature extraction framework that addresses the limitations of static 5-tuple aggregation. By jointly modeling spatial and behavioral consistency using a dual self-attention architecture and leveraging few-shot triplet learning, it enables accurate packet-flow association with minimal labeled data. Extensive evaluations show near-perfect flow identification accuracy and highly discriminative features that improve intrusion detection across heterogeneous datasets. Overall, Hi-Fi Flow offers a robust and data-efficient foundation for next-generation network monitoring and threat analytics.

REFERENCES

[1] B. Rathi, S. Thapaswi, M. Kambhampati, V. Jain, P. Akshay, T. N. Pandey, and S. K. Pradhan, "Realizing the potential of internet of things (iot) in industrial applications," *Discover Internet of Things*, vol. 5, no. 1, pp. 1–16, 2025.

[2] M. Noaman, M. S. Khan, M. F. Abrar, S. Ali, A. Alvi, and M. A. Saleem, "Challenges in integration of heterogeneous internet of things," *Scientific Programming*, vol. 2022, no. 1, p. 8626882, 2022.

[3] A. H. Lashkari, G. D. Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of tor traffic using time based features," in *International Conference on Information Systems Security and Privacy*, vol. 2. SciTePress, 2017, pp. 253–262.

[4] A. Rosay, F. Carlier, E. Cheval, and P. Leroux, "From cic-ids2017 to lycos-ids2017: A corrected dataset for better performance," in *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2021, pp. 570–575.

[5] A. Dehlaghi-Ghadim, M. H. Moghadam, A. Balador, and H. Hansson, "Anomaly detection dataset for industrial control systems," *IEEE Access*, vol. 11, pp. 107 982–107 996, 2023.

[6] M. Swarnkar, R. Kumar, R. Baidyo *et al.*, "Liteex: A lightweight feature extraction tool for captured network traces," in *2023 15th International Conference on COMmunication Systems & NETworkS (COMSNETS)*. IEEE, 2023, pp. 243–251.

[7] P. Goldschmidt and D. Chudá, "Network intrusion datasets: A survey, limitations, and recommendations," *Computers & Security*, p. 104510, 2025.

[8] B. Claise, B. Trammell, and P. Aitken, "Specification of the ip flow information export (ipfix) protocol for the exchange of flow information," Tech. Rep., 2013.

[9] P. Phaal and M. Lavine, "sFlow Version 5," sFlow.org and InMon Corp., July 2004, version 1.00, FINAL. Accessed: 2025-12-01. [Online]. Available: https://sflow.org/sflow_version_5.txt

[10] B. Claise, "Cisco systems netflow services export version 9," Tech. Rep., 2004.

[11] V. Paxson, "Bro: a system for detecting network intruders in real-time," *Computer networks*, vol. 31, no. 23-24, pp. 2435–2463, 1999.

[12] M. Mellia, A. Carpani, and R. Lo Cigno, "Tstat: Tcp statistic and analysis tool," in *International Workshop on Quality of Service in Multiservice IP Networks*. Springer, 2003, pp. 145–157.

[13] C. M. Inacio and B. Trammell, "{YAF}: Yet another flowmeter," in *24th large installation system administration conference (LISA 10)*, 2010.

[14] P. Lucente, "Pmacct: Steps forward interface counters," *Tech. Rep.*, 2008.

[15] H. C. Lee and V. L. Thing, "Port hopping for resilient networks," in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, vol. 5. IEEE, 2004, pp. 3291–3295.

[16] S. Staniford, V. Paxson, and N. Weaver, "How to own the internet in your spare time," in *11th USENIX Security Symposium (USENIX Security 02)*, 2002.

[17] B. Biggio, I. Corona, D. Maiorca, B. Nelson, N. Šrndić, P. Laskov, G. Giacinto, and F. Roli, "Evasion attacks against machine learning at test time," in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III 13*. Springer, 2013, pp. 387–402.

[18] N. Fei, Y. Gao, Z. Lu, and T. Xiang, "Z-score normalization, hubness, and few-shot learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 142–151.

[19] I. Sharafaldin, A. H. Lashkari, A. A. Ghorbani *et al.*, "Toward generating a new intrusion detection dataset and intrusion traffic characterization." *ICISSp*, vol. 1, pp. 108–116, 2018.

[20] I. Frazão, P. H. Abreu, T. Cruz, H. Araújo, and P. Simões, "Denial of service attacks: Detecting the frailties of machine learning algorithms in the classification process," in *International Conference on Critical Information Infrastructures Security*. Springer, 2018, pp. 230–235.

[21] N. Moustafa and J. Slay, "Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set)," in *2015 military communications and information systems conference (MilCIS)*. IEEE, 2015, pp. 1–6.