

Article

A Quiet Day Is Not a Secure Network: Multi-Timescale Evidence from a Single Enterprise Security Gateway

Md. Rajib Hasan ^{1,2,*}, Noor H. S. Alani ^{2,*} and Nurul I. Sarkar ¹¹ Department of Computer and Information Sciences, Auckland University of Technology, Auckland 1010, New Zealand; nurul.sarkar@aut.ac.nz² School of Computing, Eastern Institute of Technology, Napier 4142, New Zealand

* Correspondence: r.hasan@autuni.ac.nz or rhasan@eit.ac.nz (M.R.H.); nalani@eit.ac.nz (N.H.S.A.)

Abstract

Security dashboards summarise large event streams into convenient hourly, daily, weekly, and monthly views, but aggregation can change what analysts see and therefore what they prioritise. This case study examines four overlapping Check Point Quantum Spark 1530 reports from a single enterprise edge gateway: 1 h, 24 h, 7 days, and 30 days, each ending on or near 15 September 2023. We reconstruct traffic series, normalise volumes by duration, compare application concentration, reconcile summary and detailed tables, and separate control-confirmed events from unadjudicated indicators. The one-hour window corresponded to a duration-normalised 10.10 GB/day, 2.86 times the 30-day rate of 3.53 GB/day, whereas the 24 h and seven-day rates were 3.86 and 3.51 GB/day. Top five application concentration declined monotonically from 88.8% at one hour to 17.5% at 30 days. The hourly and daily summaries reported zero Anti-Bot, Anti-Virus, and intrusion-prevention events. Yet, the seven-day report contained 15,641 IPS incidents and four Anti-Virus events, and the 30-day report contained 136,057 IPS incidents, 28 Anti-Virus events, and two Anti-Bot events. Reconciliation also exposed non-additive rounded totals, partial-bin coverage, conflicting high-risk-application counts, and a weekly QUIC session count larger than its 30-day counterpart. An alignment-controlled comparison of the exactly nested 7-day/30-day pair showed nearly identical traffic rates in the final week and the preceding 23 days (3.51 versus 3.54 GB/day). In contrast, a containment bound showed that part of the apparent collapse in leading-application dominance reflects vendor reporting logic rather than aggregation. Within the limits of this single-gateway case study, the results indicate that short windows are useful for burst triage but unsafe as stand-alone evidence of security state. We propose, but do not yet validate, a multi-timescale workflow for converting vendor telemetry into cyber-risk intelligence through alignment, normalisation, reconciliation, corroboration, and explicit confidence labels.

Academic Editor: Emre Tokgoz

Received: 6 September 2026

Revised: 1 October 2026

Accepted: 3 October 2026

Published: 8 October 2026

Keywords: network telemetry; cyber-risk intelligence; temporal aggregation; security operations; intrusion prevention; dashboard consistency; risk observability; traffic analysis

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Enterprise security operations depend on telemetry. Firewalls, intrusion detection and prevention systems (IDPSs), endpoint platforms, identity systems, and cloud services continuously emit counters, alerts, flow records, and summaries. Continuous monitoring

guidance treats this evidence as a basis for maintaining awareness of threats and vulnerabilities, as well as control effectiveness [1,2]. Incident-response guidance, however, also requires analysis and contextual validation before an observation becomes an incident requiring a particular response [3]. The operational challenge, therefore, is not merely to collect more data. It is to transform heterogeneous, compressed, and sometimes inconsistent measurements into defensible cyber-risk intelligence. Foundational IDPS and log-management guidance further emphasises retaining and interpreting the underlying records [4,5].

Time scale is central to that transformation. A short window preserves bursts that a long window can smooth out, but it can also produce unstable rankings and a false impression of safety when no alert occurs. A long window improves the visibility of persistent or infrequent security events, yet aggregation and top-k reporting may suppress transient behaviour, censor smaller contributors, and mix distinct operating regimes. Temporal aggregation should therefore be treated not simply as a display preference but as a measurement operator: it changes which signals are visible, how dominant applications appear, and which control events enter the analyst's field of view. Network anomaly research has long recognised temporal and multivariate structures [6–8], while later work has developed online, deep-learning, and multi-scale detection methods [9–13]. Much of the literature, however, evaluates classifiers against labelled packet- or flow-level datasets. Practitioners often face a different artefact: a vendor-generated report containing rounded aggregates, changing category taxonomies, top-k lists, and no raw event ledger.

Source reports include traffic peaks, QUIC and LLMNR activity, uncategorised traffic, risk-labelled applications, and control-generated events that can support investigation hypotheses. These do not include packet payloads, complete flow records, alert timestamps, baselines, asset criticality, user context, change records, or analyst adjudication needed to convert those hypotheses into confirmed attacks. Treating a burst or protocol label as a unique security explanation would reproduce a closed-world error [14,15]. The analytical problem is therefore twofold: determining how the reporting horizon changes the observable security narrative, and determining whether the dashboard values are internally consistent enough to support downstream risk decisions.

We analyse four reports generated by a single Check Point Quantum Spark 1530 appliance. The reports cover approximately one hour, 24 h, seven days, and 30 days, ending between 16:00 and 17:51 on 15 September 2023. They describe one gateway and therefore constitute repeated, differently aggregated views of one operational environment rather than four independent samples. Importantly, the 7-day and 30-day reports end at the same time and form a genuinely nested pair, permitting consistency checks that are not defensible for the offset 1 h and 24 h windows. The one-hour window, by contrast, begins 51 min after the 7-day/30-day pair ends and overlaps with the daily window by only nine minutes, so comparisons involving it combine a reporting-horizon effect with a time-of-observation effect; Section 4.7 separates these components where the data allow.

The study addresses three research questions:

- RQ1: How does reporting horizon reshape traffic intensity, burst structure, application concentration, and the stability of the top-10 application set?
- RQ2: How does reporting horizon affect the emergence and temporal localisation of control-confirmed events and risk-bearing activity?
- RQ3: Which arithmetic, semantic, and nested-window consistency constraints fail across reports, and what do those failures imply for confidence in dashboard-derived cyber-risk intelligence?

The paper makes five contributions. First, it provides a de-identified single-gateway case study of commercial gateway telemetry across four reporting horizons. Second, it develops a multiscale measurement framework combining duration normalisation,

within-window intensity, burst metrics, cumulative concentration, and application-set similarity. Third, it introduces an exposure-adjusted localisation analysis that compares activity in the final week with activity accumulated earlier in the 30-day window, relative to uniform and weekly-periodic exposure references and with a clustering-aware sensitivity analysis. Fourth, it treats reconciliation failures as measurable evidence-quality signals through nested-window monotonicity tests, containment bounds, and signed dashboard residuals, and uses alignment-controlled comparisons to separate aggregation effects from time-of-observation effects and vendor reporting logic. Fifth, it proposes, as a structured analytic convention that still requires independent validation, a confidence-aware operational workflow that grades observations, telemetry indicators, control-confirmed events, and analyst-confirmed incidents on an ordinal scale of evidential commitment with explicit transition rules, and records explanatory hypotheses separately as a non-ordinal annotation.

2. Background and Related Work

2.1. Network Telemetry and Intrusion Detection

Network monitoring has evolved from rule-driven inspection at a strategic point to a distributed analysis of flows, sessions, applications, and host activity. Bro demonstrated how policy-oriented real-time analysis could connect packet observations to higher-level events [6]. Network-wide traffic studies showed that aggregate measurements can reveal anomalies when analysed across links and dimensions [7]. Anomaly-based intrusion detection subsequently developed statistical, knowledge-based, and machine-learning approaches to distinguish normal behaviour from departures that may indicate misuse [8,16].

Anomaly detection is most useful when signatures are incomplete, or attacks are novel. At the same time, anomaly detectors can generate false positives because workload changes, updates, backups, congestion, and user behaviour also produce deviations. NIST distinguishes detection and prevention capabilities from incident handling and recommends analysing security data in context [3,4]. This distinction is especially important for appliance reports. A traffic peak is an observation; high QUIC or LLMNR activity is an indicator that may merit investigation; a control-generated malware or IPS event is stronger evidence; and a confirmed incident requires corroboration and adjudication.

Machine-learning research has expanded the detection space. Kitsune used an ensemble of autoencoders for efficient online network intrusion detection [9], and researchers have applied deep representations to non-linear traffic features [10]. However, deployment validity remains difficult. Sommer and Paxson argued that security monitoring operates outside a closed world, with evolving traffic and a severe base-rate problem [14]. Arp et al. catalogued experimental pitfalls that inflate security-ML performance [15], while TESSERACT demonstrated how temporal and spatial bias can make evaluation results unrealistic [17]. These cautions support a conservative interpretation of unlabelled enterprise telemetry.

2.2. Datasets, Aggregation, and Temporal Scale

Public intrusion datasets enable reproducible benchmarking but differ substantially in age, attack coverage, network topology, feature extraction, and realism [18–20]. Newer datasets such as Bot-IoT were designed to represent more contemporary and heterogeneous traffic [21], yet any benchmark remains a model of deployment rather than the deployment itself. Host-based evidence can complement network data, but it introduces its own collection and interpretation constraints [22]. For an operational gateway, validation

therefore requires correlation with evidence from endpoints, identity, DNS, vulnerability, and change management rather than reliance on a single traffic view.

Aggregation affects both detection and interpretation. Lakhina et al. showed that correlated measurements can diagnose network-wide anomalies [7]. Recent methods explicitly learn across multiple observation or transformation scales. Duan et al. used multi-scale windows and time–frequency features to improve anomaly classification [11]. TTANAD used temporal test-time augmentation to provide multiple perspectives at inference [12]. Liu et al. combined multi-scale subspace analysis with symbolic learning for intrusion detection [13]. These methods treat scale as a modelling dimension. The present study treats it as an evidence-governance dimension: even before training a model, analysts must understand how the reporting horizon changes totals, rankings, and event visibility. In spatio-temporal analysis, the dependence of results on the chosen temporal unit is known as the modifiable temporal unit problem [23]; dashboard reporting horizons are an operational instance of that problem.

Log-anomaly research likewise shows that sequence context, feature construction, and data quality strongly shape the results [24]. Vendor dashboards add a layer because they typically expose rounded aggregates, ranked subsets, and summary KPIs whose definitions may change across views. Such interfaces are useful for rapid triage, but researchers rarely examine their measurement validity directly. In particular, a value can be operationally plausible while still being non-additive, top-k censored, semantically incompatible with another panel, or impossible under a nested-window count constraint. These properties matter before any anomaly model, fusion algorithm, or executive risk score is applied.

2.3. From Threat Information to Cyber-Risk Intelligence

Threat information becomes intelligence when it is evaluated, contextualised, and connected to decisions. The technical cyber-threat intelligence literature emphasises evidence about adversaries, indicators, and defensive actions [25]. Organisational research further argues that CTI must synthesise information into stakeholder- and decision-relevant products, rather than merely increasing the volume of technical data [26,27]. Recent work has examined multi-granular fusion [28], knowledge-based entity and relation extraction [29], and privacy-preserving correlation across organisations [30].

The present study adopts that decision orientation but focuses on locally generated operational telemetry. The objective is not to label every burst as malicious. Instead, it determines which observations remain stable across horizons, which emerge only after temporal aggregation, which conflict with other dashboard fields, what additional evidence is required, and which operational decision is justified at each confidence level. In this sense, cyber-risk intelligence is treated as a confidence-bearing decision product rather than a larger collection of threat-labelled counts. This aligns continuous monitoring with the Govern, Identify, Protect, Detect, Respond, and Recover functions of the NIST Cybersecurity Framework 2.0 [2].

The proposed workflow overlaps with established data-quality and telemetry-validation frameworks, and its contribution is best stated relative to them. General data-quality research defines intrinsic, contextual, representational, and accessibility dimensions such as accuracy, completeness, consistency, and timeliness [31]. Frameworks specific to cyber-threat intelligence operationalise comparable dimensions for shared threat artefacts and support automated and expert-driven quality scoring [32,33]. At the same time, studies of security operations centres document how alert volume, false positives, and missing context degrade analyst decisions [34,35]. These frameworks assess the quality of an artefact or feed at a given point in time. They do not specify how the same metric should behave when one reporting window is contained in another, nor do they separate the

effect of the aggregation horizon from the time at which a window is observed. The present workflow therefore adds three scale-specific checks to these established dimensions: nested-window monotonicity and containment bounds, alignment-controlled comparison of exactly nested windows, and exposure-adjusted localisation. Section 4.7 reports the measurable added value of these checks in the present data, namely the number of material inconsistencies that single-report validation would not detect.

QUIC illustrates the need for context. QUIC is a standardised encrypted transport over UDP, designed for secure, multiplexed connections and reduced latency [36]. A high QUIC volume or session count is not inherently malicious. It can reflect ordinary web and application traffic, but encryption and rapid adoption can challenge legacy visibility. Risk arises from unexpected behaviour relative to policy and baseline, not from the protocol name alone.

3. Materials and Methods

3.1. Study Design and Unit of Analysis

We conducted a retrospective descriptive case study of a single enterprise edge gateway. The primary sources were four Check Point Quantum Spark 1530 appliance installed at BICC Cyber security & IT Solutions at 131, Blockhouse bay Road, Auckland 1023, New Zealand reports, and the unit of analysis is the report window. Because the windows substantially overlap and refer to the same gateway, they are not independent replicates. The analysis therefore uses descriptive comparisons and internal consistency checks; it does not use null-hypothesis significance tests or estimate population-level attack rates.

3.2. Report Windows and Source Characteristics

Table 1 summarises the sources. The so-called monthly report is a 30-day window from 16 August to 15 September, not a calendar-month September report. The one-hour and daily reports end later than the weekly and 30-day reports. This offset prevents exact set containment and must be considered when comparing the final hour or day. Table 1 implies the following overlaps. The one-hour window lies entirely after the 7-day and 30-day windows and overlaps the daily window by only nine minutes; the daily window overlaps the 7-day and 30-day windows for 23 of its 24 h; and only the 7-day window is exactly contained in the 30-day window. Comparisons involving the one-hour and daily reports therefore confound reporting horizon with observation time and are interpreted descriptively, whereas horizon-attribution claims are restricted to the aligned 7-day/30-day pair (Section 4.7).

Table 1. Overlapping report windows analysed in the study. Times are reproduced from report headers. The gateway identifier, MAC address, internal host addresses, and public site addresses are omitted or pseudonymized.

Reporting Horizon h	Source Interval	Duration D_h (Hours)	Native Bin Width Δ_h	Total Traffic V_h (GB)	Security-Event Variables Available
Hourly	15 September 2023, 16:51–17:51	1	5 min	0.421	Anti-Bot, Anti-Virus, IPS, threat emulation, high-risk applications
Daily	14 September 2023, 17:00–15 September 2023, 17:00	24	1 h	3.86	Same categories and high-risk applications
Weekly	8 September 2023, 16:00–15 September 2023, 16:00	168	12 h	24.6	Same categories, protection names, malware records and host records

30-day	16 August 2023, 16:00–15 September 2023, 16:00	720	24 h	106	Same categories, protection names, malware records and host records
--------	--	-----	------	-----	--

3.3. Extraction, De-Identification, and Verification

The Check Point Quantum Spark security appliance generated the network and security data directly through its built-in reporting functionality. For each reporting period, the extracted variables included reporting-window timestamps; total, received, and sent network traffic; temporal traffic measurements; leading applications and sites by bandwidth and session count; traffic categories; security-control event totals; intrusion-prevention events and protections; malware detections; potentially high-risk applications; and host-level security incidents. Values were recorded at the precision the appliance provided. We retained traffic volumes reported in MB and GB in their original units, without converting to bytes, because the reporting output did not specify whether these units used binary or decimal conventions.

We independently cross-checked all values used in the analysis and graphical presentation against the corresponding Check Point appliance-generated reports to minimise transcription and data-entry errors. We then de-identified the dataset before analysis and manuscript preparation. We removed infrastructure identifiers that could disclose the monitored network environment. When comparative analysis required individual public IP addresses, we pseudonymised them as Site A, Site B, and Site C, and replaced private host addresses with identifiers such as Host H1–H4. We retained application, security-product, protocol, malware, and protection names when needed for technical interpretation, as they represented analytically relevant cybersecurity attributes rather than identifying information.

We treated the appliance-generated data as observational network-security telemetry, and restricted interpretations to information directly supported by the recorded traffic and security events. In particular, detection events, intrusion-prevention incidents, potentially high-risk application activity, and host-level incidents were not automatically interpreted as evidence of successful compromise unless the available security records supported such a conclusion. Technical descriptions of the monitoring environment were similarly limited to characteristics documented by the available Check Point appliance records. Accordingly, the monitored gateway is described as an enterprise network-edge security appliance, without assuming a public-cloud architecture, virtual-network deployment, tenancy model, or cloud control plane, as none were documented in the available system information.

3.4. Measures

Table 1 defines the dataset variables used in the analysis. Each reporting horizon is indexed by h , where $h \in \{\text{hourly, daily, weekly, 30 – day}\}$. For each horizon, D_h denotes the duration of the reporting window in hours, V_h denotes the total traffic reported during that window in gigabytes, and Δ_h denotes the native temporal bin width. Because the four reporting windows have different durations, their unadjusted traffic volumes are not directly comparable. Total traffic was therefore converted to a duration-normalised rate expressed on a per-day scale:

$$R_h = \frac{V_h}{D_h/24} = \frac{24V_h}{D_h} \quad (1)$$

where R_h is the duration-normalised traffic rate for horizon h , expressed in GB day^{−1}; V_h is the total traffic volume in GB obtained from Table 1; and D_h is the corresponding reporting-window duration in hours. Thus, every variable in Equation (1) is directly linked to an observed or derived field in Table 1. The source intervals and security-event

coverage fields in Table 1 are not mathematical inputs to Equation (1). We used the source intervals to assess window alignment and containment, while security-event coverage determined which control-event measures we could compare across horizons.

For example, the hourly report contained $V_{\text{hourly}} = 0.421$ GB over $D_{\text{hourly}} = 1$ h. Therefore, its duration-normalised rate was

$$R_{\text{hourly}} = \frac{24(0.421)}{1} = 10.10 \text{ GB day}^{-1}$$

Applying the same calculation produced duration-normalised rates of 3.86, 3.51, and 3.53 GB day^{−1} for the daily, weekly, and 30-day horizons, respectively.

Because R_h rescales the traffic of one window to a per-day scale, it is a duration-normalised quantity rather than an estimate of the traffic that would have been observed over a full day; for the one-hour window it rescales a single interval of a bursty process. We therefore also report the matched hourly rate, V_h/D_h (0.421, 0.161, 0.146, and 0.147 GB h^{−1} for the hourly, daily, weekly, and 30-day reports) and locate the one-hour and daily values within the empirical distributions of equal-length bins available in the daily and 30-day reports (Section 4.7).

Second, application concentration was measured as

$$C_{k,h} = \frac{V_{\text{top-}k,h}}{V_h} \quad (2)$$

where $V_{\text{top-}k,h}$ is the traffic attributed to the k highest-volume applications within horizon h , and V_h is the corresponding total traffic volume reported in Table 1. We calculated values for $k = 5$ and $k = 10$.

Third, the native bin width Δ_h in Table 1 determines the temporal observations available for each reporting horizon. Let $V_{t,h}$ denote the traffic volume observed in temporal bin t within horizon h . Within-report burstiness was measured using

$$B_h = \frac{\max_t(V_{t,h})}{\text{median}_t(V_{t,h})} \quad (3)$$

and

$$CV_h = \frac{\sigma_t(V_{t,h})}{\text{mean}_t(V_{t,h})} \quad (4)$$

where $\sigma_t(\cdot)$ and $\text{mean}_t(\cdot)$ denote the standard deviation and arithmetic mean of $V_{t,h}$ over the native bins t of horizon h .

Because Δ_h differs across the four horizons—from five-minute bins to daily bins—these measures describe within-report temporal variation and are not treated as directly equivalent inferential statistics. To examine whether the cross-horizon ordering of burstiness depends on bin width, we additionally computed a resolution-harmonised coefficient of variation, $CV_h^* = CV_h \cdot \sqrt{\Delta_h/24 \text{ h}}$, which rescales each coefficient to the value it would take at a common 24 h resolution if bins were independent and identically distributed. Because the four coefficients are computed from different, only partly overlapping observation windows, CV^* cannot separate the effect of bin width from differences in the workload observed in each window, including non-stationarity within and between windows [37]. CV^* is therefore used only as a diagnostic of the sensitivity of the burst metrics to temporal resolution, not as evidence of temporal dependence (Section 4.7).

We measured top-10 membership stability between horizons with the Jaccard index, $J(A, B) = |A \cap B| / |A \cup B|$ of the displayed application-name sets; vendor labels were not semantically merged beyond trivial formatting harmonisation.

Fourth, we assessed security visibility using Anti-Bot events, Anti-Virus events, IPS incidents, listed potentially high-risk application sessions, and detailed malware/protection/host rows. We calculated reconciliation deltas between cover KPIs and detailed row sums. We tested monotonic constraints that should hold when a shorter interval is

contained within a longer interval under stable definitions. For the exactly nested pair, we defined the exposure-adjusted localisation ratio as $L_x = \log_2[(X_{7d}/X_{30d})/(7/30)]$, where X is any cumulative count or volume reported in both windows; zero denotes accumulation proportional to elapsed time, positive values concentration in the final week, and negative values concentration earlier in the 30 days. Because the reference 7/30 assumes temporally uniform exposure, L_x measures deviation from uniform exposure rather than evidence of recent abnormality; Section 4.7 examines its sensitivity to weekday, weekend, and business-hour structure and to temporal clustering of events. Second, a containment lower bound on application concentration was derived: under a stable taxonomy, the k leading 30-day applications cannot carry less traffic than the k leading 7-day applications, so $C_{k,30d} \geq C_{k,7d} \cdot V_{7d}/V_{30d}$.

3.5. Evidence and Confidence Model

To prevent telemetry from being mistaken for incident attribution, findings were graded on an ordinal scale of evidential commitment, and candidate explanations were recorded separately as hypotheses:

- Level A—Control-Confirmed Event: A named security control recorded an event or incident, such as an Anti-Virus event, an IPS incident, or a malware row. This confirms control activity, not necessarily a successful compromise.
- Level B—Telemetry Indicator: A measured deviation or exposure-relevant feature, such as a traffic burst, inbound/outbound imbalance, high session count, uncategorized traffic, or use of a high-risk application. The indicator justifies triage but does not identify an attack.
- Explanatory Hypothesis (Non-Ordinal Annotation): A possible benign or malicious mechanism, such as an update, backup, scanning, DDoS, exfiltration, spoofing, or command-and-control. A hypothesis can be attached to a finding at any evidence level; it is not itself an evidence level and does not raise the level of the finding it explains. A hypothesis requires corroboration before operational escalation beyond triage.

This model aligns with the continuous-monitoring and incident-response guidance, which requires interpreting security observations in relation to assets, threats, control performance, and organisational risk [1–4]. Two boundary states complete the ordinal scale used in Section 5.3, giving four ordered states: observation (a reported value before interpretation) < Level B telemetry indicator < Level A control-confirmed event < analyst-confirmed incident. The last state requires adjudication that the source reports cannot provide, so no finding in this study reaches it. The model is proposed as an analytic convention; its inter-analyst reliability has not yet been tested (Section 5.6).

The ordering does not rank how serious or how plausible an interpretation is. It ranks evidential commitment: each higher state requires that one further, independent party, namely a reconciliation check, a named security control, or a human adjudicator, has classified the activity. Explanatory hypotheses are excluded from the scale for this reason. A hypothesis adds interpretation rather than evidence, and it can be attached to an observation as readily as to a control-confirmed event, so placing it between an indicator and a control-confirmed event would imply an evidential gain that has not occurred. Transitions between states follow four rules: (i) Observation to Level B: The reported value passes the reconciliation checks in Section 3.4 and either deviates from a stated reference, such as a matched baseline or exposure reference, or matches a pre-specified exposure-relevant feature. (ii) Entry at Level A: A named security control records an event attributable to the same activity; a finding may enter Level A directly, without first being a Level B indicator, because the control itself performs the classification. (iii) Level B or Level A to Analyst-

Confirmed Incident: An analyst adjudicates the finding using at least one evidence source independent of the gateway, such as endpoint, identity, DNS, or change-management records [3]. (iv) Demotion: A finding whose supporting value fails a reconciliation check is held at the observation state and flagged as a data-quality issue until the discrepancy is resolved, and disconfirming evidence returns a finding to the highest state that the remaining evidence supports. Hypotheses may be added, retained, or rejected at any state, and confirming a hypothesis changes the state only through rule (iii). Because the states are ordered by the number and independence of the parties that have classified the activity, rather than by a fixed sequence of analytic steps, the scale remains ordinal when a finding skips a state.

3.6. Analysis Procedure

The analysis followed nine steps: (1) align report windows and identify offsets; (2) normalise traffic by duration; (3) reconstruct report-native profiles and compute within-window intensity and burst metrics; (4) calculate cumulative application concentration; (5) compare top-10 application membership using Jaccard similarity; (6) quantify security-evidence emergence and exposure-adjusted localisation; (7) audit nested-window monotonicity, containment bounds, and dashboard residuals; (8) perform alignment-controlled and resolution-harmonised sensitivity analyses; and (9) translate the resulting evidence into confidence-graded risk questions and validation actions. All analyses are descriptive and were implemented in R (version 4.6.1 windows) using reproducible scripts; no model was trained because the reports provide neither raw observations nor labelled ground truth.

4. Results

4.1. Traffic Rate and Direction Change with Reporting Horizon

The reports recorded 421 MB in one hour, 3.86 GB in 24 h, 24.6 GB in seven days, and 106 GB in 30 days (Table 2). After duration normalisation, the one-hour interval averaged 10.10 GB/day, 2.86 times the 30-day average of 3.53 GB/day. The daily rate was 3.86 GB/day, 1.09 times the 30-day average. Weekly and 30-day rates were nearly identical at 3.51 and 3.53 GB/day. The final observed hour was therefore a high-activity interval relative to the longer-horizon mean rate, whereas the final day was only moderately elevated; Section 4.7 shows, however, that both values lie within the ordinary variation in equal-length intervals. Received traffic dominated at every horizon. Component-based received shares ranged from 69.2% in the weekly report to 78.2% in the hourly report (see Figure 1). The proportion became more inbound-heavy in the short window, but direction alone does not distinguish a benign download, update, synchronisation task, user workload, or hostile inbound activity.

Table 2. Traffic and application concentration by horizon. Received share uses received/(received + sent) to reduce distortion from rounded cover totals.

Horizon	Reported Total	Received	Sent	Equivalent GB/Day	Received Share	Top-Five Share	Top-Ten Share
Hourly	0.421 GB	0.331 GB	0.0923 GB	10.10	78.2%	88.8%	96%
Daily	3.86 GB	2.79 GB	1.08 GB	3.86	72.1%	77.2%	83%
Weekly	24.6 GB	17.1 GB	7.61 GB	3.51	69.2%	61.8%	79%
30-day	106 GB	75.5 GB	31.5 GB	3.53	70.6%	17.5%	25%

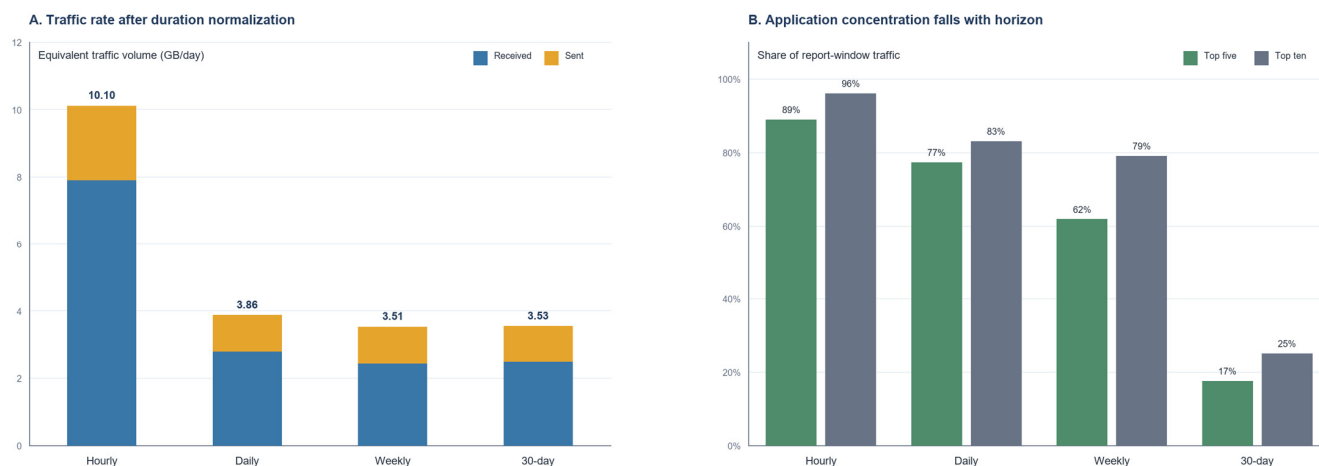


Figure 1. Duration-normalised traffic and application concentration by reporting horizon. (A) Duration-normalised traffic rate. (B) Top-five and top-ten application shares. Windows overlap but are not perfectly aligned.

4.2. Horizon-Dependent Burst Structure and Within-Window Intensity

Figure 2A expresses each report-native traffic bin as a multiple of that window's median, making the internal shape of each horizon comparable without pretending that their bin widths are equivalent. The one-hour profile contained two clear short-lived bursts, reaching approximately 4.6 times the within-hour median. The 24 h profile was substantially more episodic: the late-window business-day peaks reached approximately 25 times the daily median, whereas most overnight bins remained near or below the median. The 7-day and 30-day profiles showed recurrent elevated periods but much lower relative amplitudes, consistent with longer aggregation smoothing short-lived extremes. Because native bin widths differ, these profiles and the associated burst metrics characterise only within-window structure.

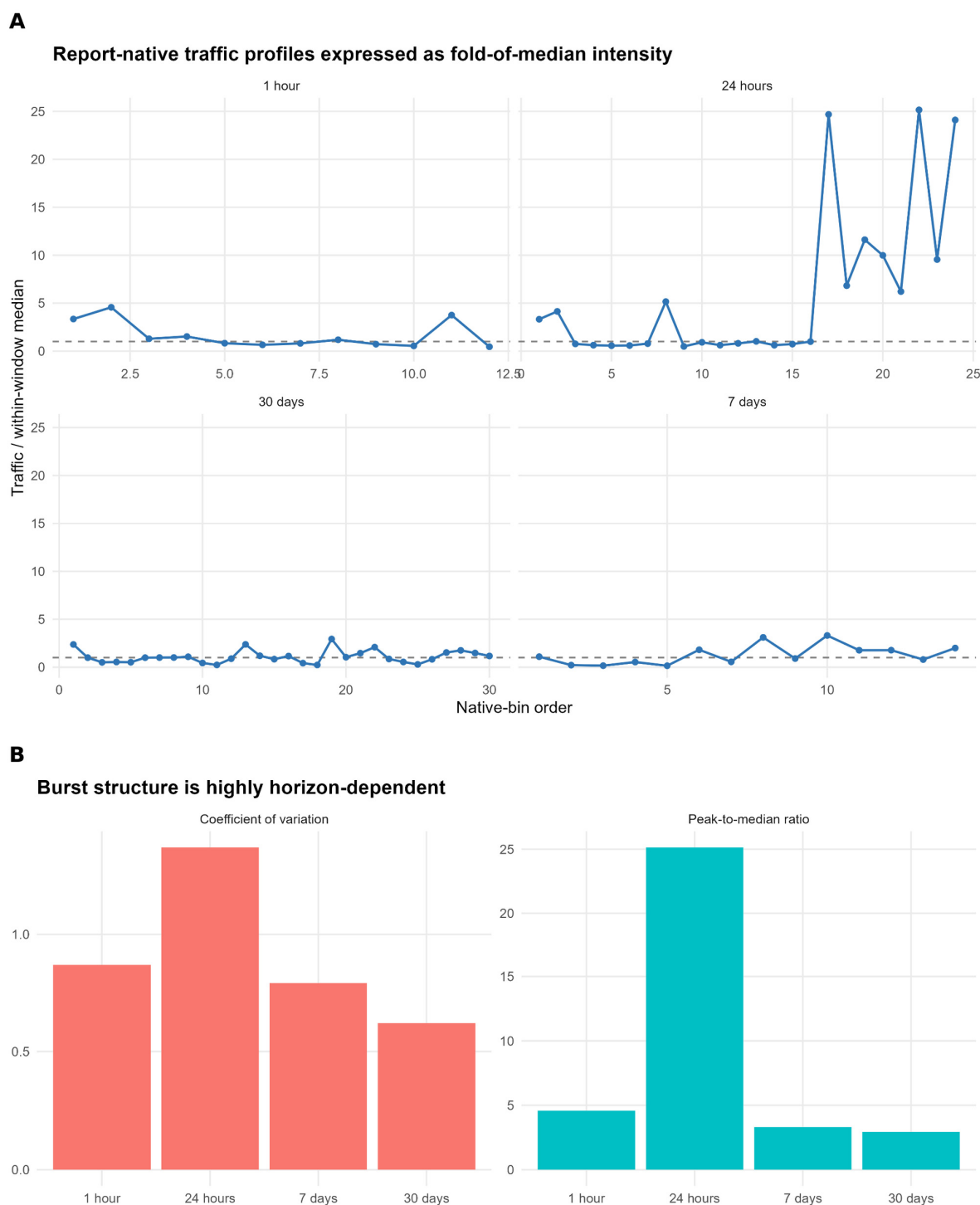


Figure 2. Horizon-dependent burst structure. (A) Report-native traffic as fold-of-median intensity; x -axes show native-bin order and are not time-aligned. (B) Coefficient of variation and peak-to-median ratio.

The formal burst metrics support the same interpretation (Figure 2B). The 24 h window had the highest coefficient of variation (1.37) and peak-to-median ratio (25.14), compared with 0.87 and 4.57 at one hour, 0.79 and 3.31 at seven days, and 0.62 and 2.93 at 30 days. Thus, the most extreme relative burstiness was not observed in the shortest report but in the 24 h horizon, where a quiet overnight baseline coexisted with several large day-time peaks.

Absolute peaks nevertheless remained operationally relevant. The weekly series reached approximately 4.47–4.48 GB in the 12 h bin labelled 13 September 17:00, and the

30-day series reached approximately 9.50–9.55 GB on 4 September. Because the bin widths differ, these values should not be compared as estimates of a single variance process. Their value is evidential: they show that each horizon reveals a different temporal structure, with short windows preserving burst timing and longer windows exposing the recurrence of high-volume regimes.

4.3. Application Concentration Declines and Top-10 Membership Is Unstable

Application concentration declined sharply as the horizon expanded (Figure 3A). The leading application alone accounted for 70% of hourly traffic, 41% of daily traffic, and 30% of weekly traffic, but only 4.9% of 30-day traffic. The cumulative top-10 shares were 96%, 83%, 79%, and 25%, respectively. The divergence is therefore not confined to an arbitrary top-five threshold: the entire rank–concentration curve flattens with longer observation, showing that short-window dashboards can create a much narrower workload narrative than the longer-run environment.

The identity of prominent applications also changed with horizon. QUIC (RFC 9000) dominated bandwidth in the hourly, daily, and weekly views, whereas Windows 11 update was the largest named application in the 30-day table. This change is compatible with ordinary workload mixing and vendor taxonomy behaviour; it does not imply that QUIC became safer or more dangerous as the horizon expanded.

Top-10 set similarity makes this reprioritisation explicit (Figure 3B). Jaccard similarity ranged from 0.25 to 0.82. The 24 h and 7-day application sets were most similar ($J = 0.82$), suggesting relatively stable membership across adjacent operational horizons, while the 1 h and 30-day sets were least similar ($J = 0.25$). The remaining pairwise similarities were 0.33 or 0.43. Importantly, adjacent windows can retain similar application membership while assigning very different traffic shares to those applications; horizon affects both who appears in the top list and how dominant each member appears. High Jaccard values indicate stable membership, and low values indicate horizon-dependent application churn.

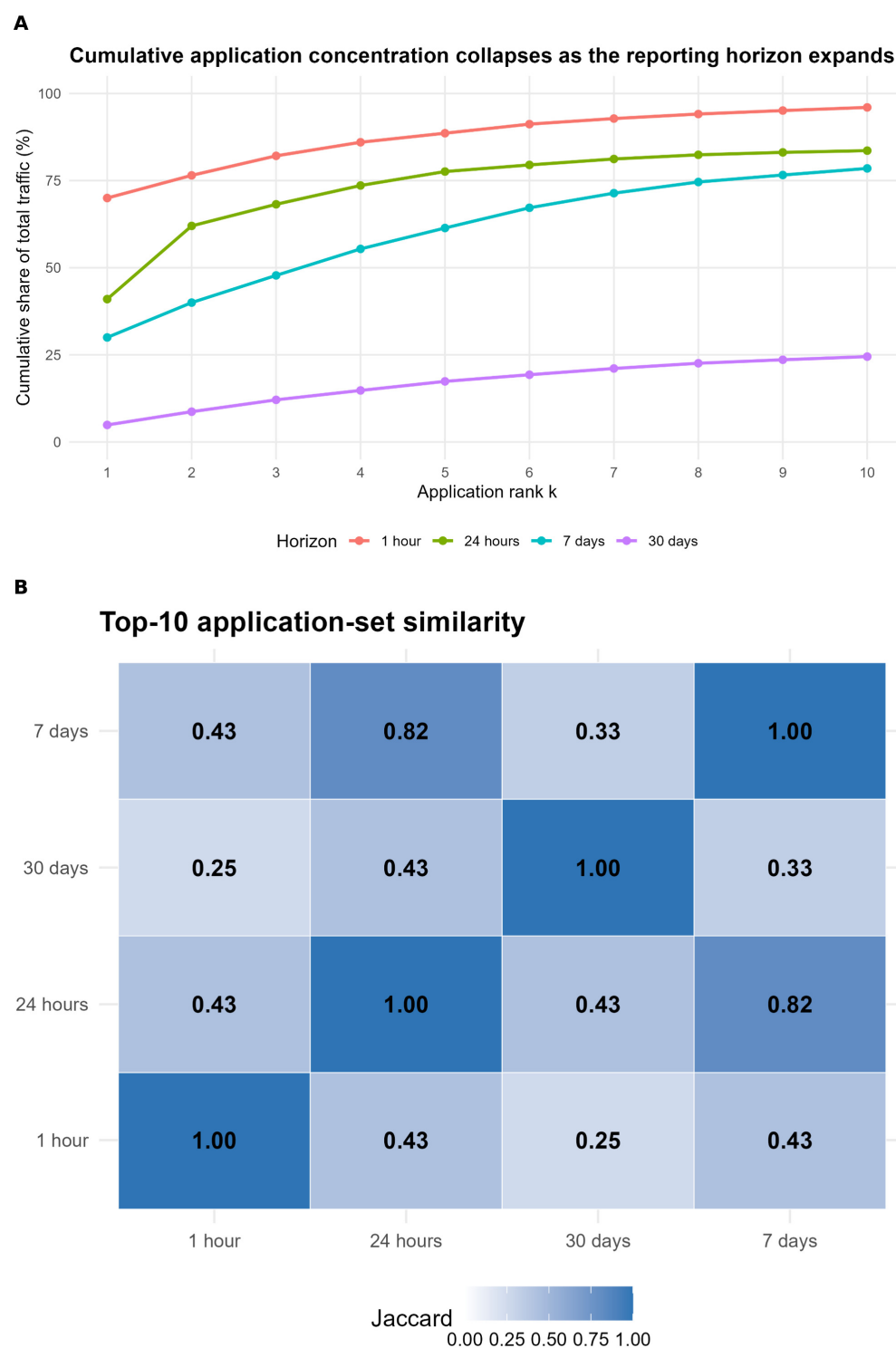


Figure 3. Horizon-dependent application concentration and set stability. (A) Cumulative traffic share of the first k displayed applications ($k = 1, \dots, 10$). (B) Jaccard similarity of top-10 application-name sets; darker cells denote higher similarity.

4.4. Security Evidence Emerges Nonlinearly and Is Temporally Localised

The strongest horizon effect concerned the visibility of security evidence (Figure 4A). The one-hour and 24 h cover summaries reported zero Anti-Bot events, zero Anti-Virus events, and zero IPS incidents. However, Table 3 depicts that the 7-day report recorded four Anti-Virus events and 15,641 IPS incidents, and the 30-day report recorded two Anti-Bot events, 28 Anti-Virus events, and 136,057 IPS incidents. The detailed high-risk application totals simultaneously increased from 1 session at one hour to 17 at 24 h, 99 at seven

days, and 920 at 30 days. Malware and host-level incident detail appeared only in the longer reports, with two malware rows/two hosts at seven days and ten malware incidents/four hosts at 30 days. The matrix should therefore be read as an evidence-visibility map rather than a common-scale event-rate comparison; fine-window zeros do not make unavailable historical detail equivalent to confirmed absence.

Because the 7-day interval is nested within the 30-day interval, we can also assess the relative temporal location of shared activity (Figure 4B). The week represented 23.3% of the 30-day exposure. Total traffic (23.2%), received traffic (22.6%), and sent traffic (24.2%) were close to proportional accumulation, placing their exposure-adjusted localisation ratios near zero relative to a uniform-exposure reference. Security-control activity was less concentrated in the final week: the week contained 11.5% of 30-day IPS incidents and 14.3% of Anti-Virus events. Detailed high-risk sessions (10.8%), TeamViewer sessions (15.4%), and WireGuard sessions (7.6%) were also relatively sparse in the final week.

Other indicators were disproportionately localised in the final week. Uncategorised traffic accounted for 31.3% of its 30-day volume and approximately 30.6% of its 30-day category sessions, while two of four hosts in the 30-day incident detail appeared in the weekly detail. TunnelBear was listed four times in both the weekly and 30-day reports, which places the displayed activity entirely in the final week if the metric semantics are stable. These ratios describe deviation from uniform exposure, not severity or recent abnormality: treat large values from small counts as triage cues rather than evidence of elevated threat.

The longer-window control detail also showed repeated network enforcement rather than a single isolated event. TCP Segment Limit Enforcement dominated both the 7-day and 30-day protection tables (rounded 16 K and 136 K sessions, respectively), followed by much smaller counts for Max Ping Size and TCP Urgent Data Enforcement. These values confirm that named controls were repeatedly triggered, but vendor incident/session semantics and deduplication rules were unavailable; they should not be interpreted as numbers of distinct attacker actions or successful compromises.

Table 3. Security-control visibility and high-risk-application reconciliation. The cover KPI label does not specify whether it is a session, event, or distinct-application count; therefore, it is reported separately from detailed-row sums.

Horizon	Anti-Bot Events	Anti-Virus Events	IPS Incidents	Cover High-Risk KPI	Sum of Listed High-Risk Sessions
Hourly	0	0	0	1	1
Daily	0	0	0	16	17
Weekly	0	4	15,641	108	99
30-day	2	28	136,057	833	920

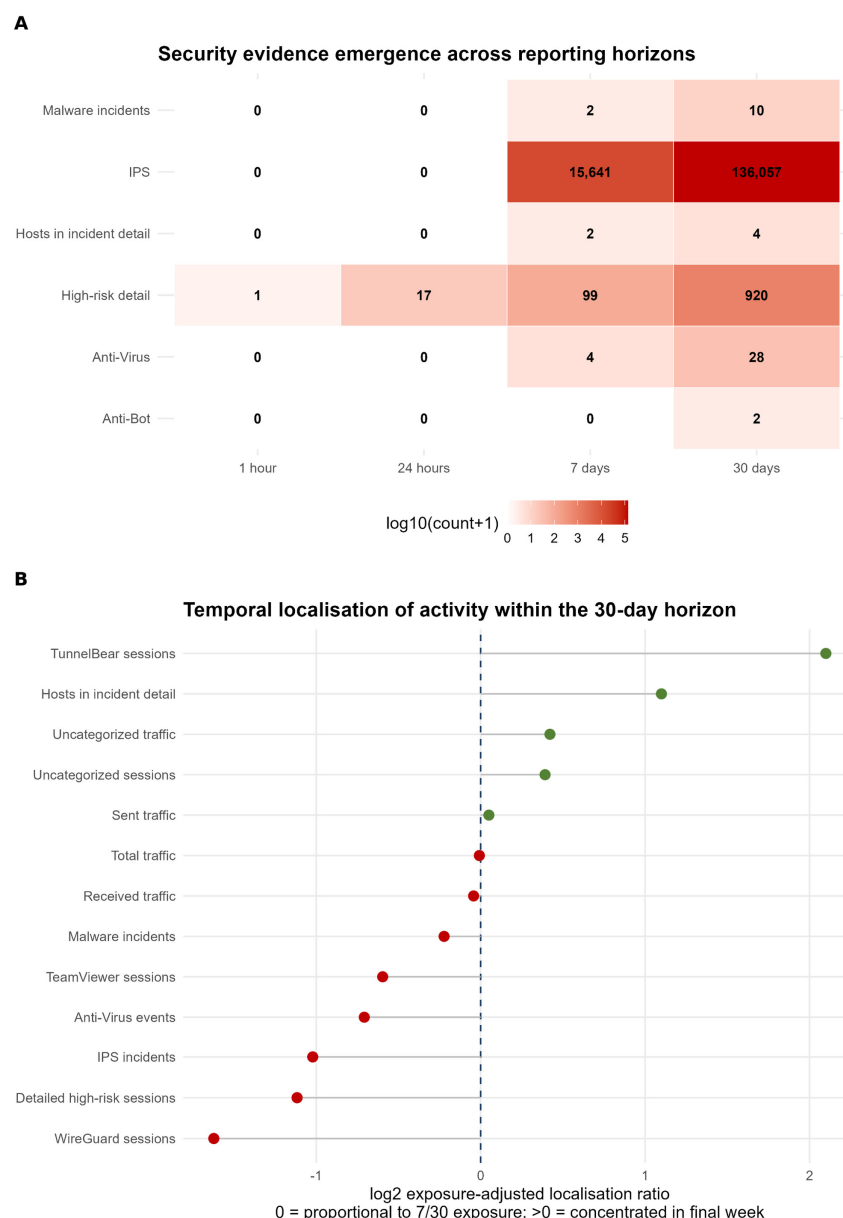


Figure 4. Security evidence emergence and temporal localisation. **(A)** Reported security evidence by horizon: fill uses $\log_{10}(\text{count} + 1)$; labels show reported counts. **(B)** Exposure-adjusted localisation within the nested 7-day/30-day windows (Section 3.4).

4.5. Risk-Bearing Application Visibility and Classification Uncertainty Expand with Horizon

Risk-labelled application visibility expanded materially with horizon (Figure 5A). TeamViewer was the only displayed high-risk application in the one-hour and 24 h detail, with 1 and 17 sessions. The 7-day view added WireGuard (5) and TunnelBear (4), while the 30-day view added BitTorrent (241), WireGuard (66), LogMeIn Rescue (14), LastPass (7), TunnelBear (4), TeamViewer Web Management (3), and OpenVPN (2), alongside 583 TeamViewer sessions. The figure deliberately distinguishes “Not listed” from zero: vendor top-k omission does not establish that an application was absent from the underlying traffic. The security question is therefore whether displayed use is authorised, attributable, and consistent with the asset role, not whether the product name itself proves maliciousness.

Classification uncertainty showed a similarly strong horizon effect (Figure 5B). Uncategorised traffic increased from 1.18 MB in the one-hour report to 450 MB in the 24 h report, 3.5 GB in the 7-day report, and 11.2 GB in the 30-day report. More importantly, its

share of total traffic rose from only 0.28% in the one-hour view to approximately 12%, 14%, and 11% in the daily, weekly, and 30-day views. The longer horizons therefore reveal a persistent 11–14% classification gap rather than merely an accumulation of bytes. Unknown or uncategorised traffic is not synonymous with malware, but persistent inability to identify application semantics weakens policy enforcement, attribution, and confidence in behavioural baselines. The persistence of this gap at the 7-day and 30-day horizons shows that it is not merely an effect of window duration.

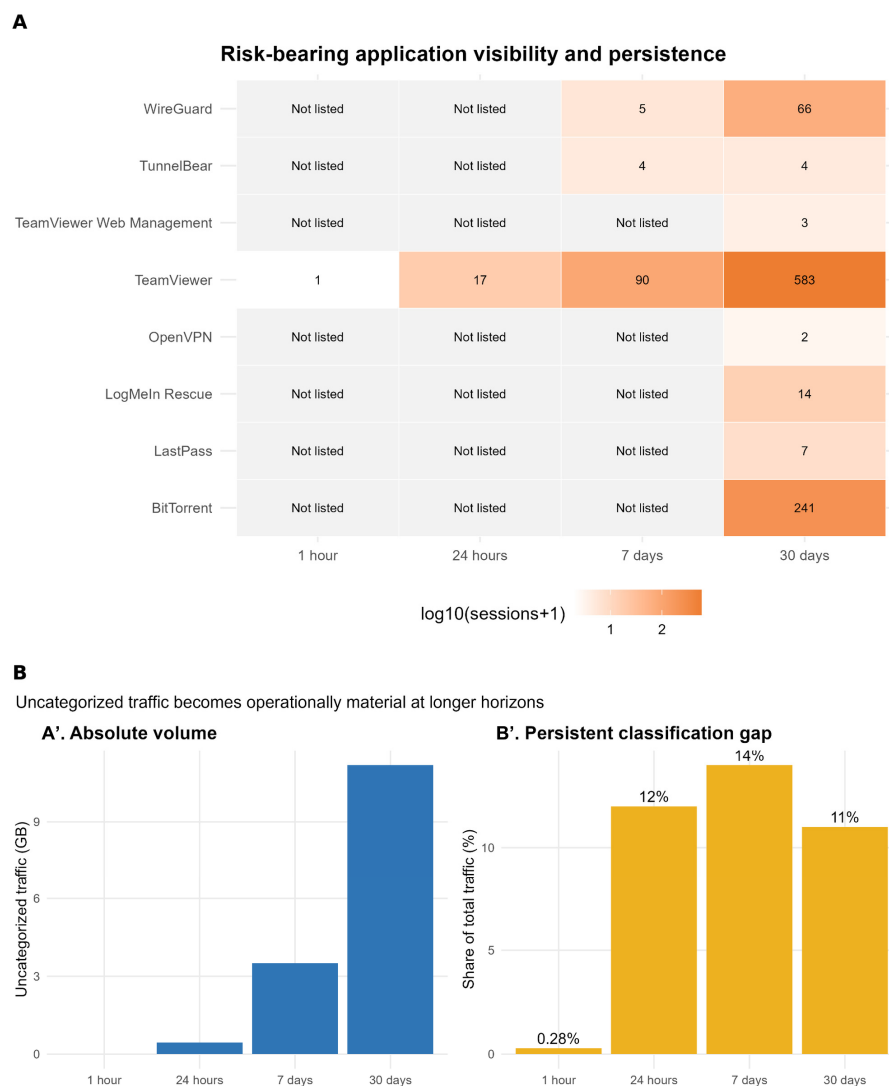


Figure 5. Risk-bearing application visibility and classification uncertainty. (A) High-risk applications displayed by the vendor; “Not listed” denotes a censored state, not zero. (B) Uncategorized traffic volume and share of total traffic.

4.6. Nested-Window Monotonicity and Reconciliation Expose Measurement Limits

The final analysis asks whether the dashboard values are mutually compatible before using them as cyber-risk evidence. Five material reconciliation issues were identified in the source reports (Table 4), and the two quantitative diagnostics in Figure 6 show that the inconsistencies occur at both cross-horizon and within-report levels.

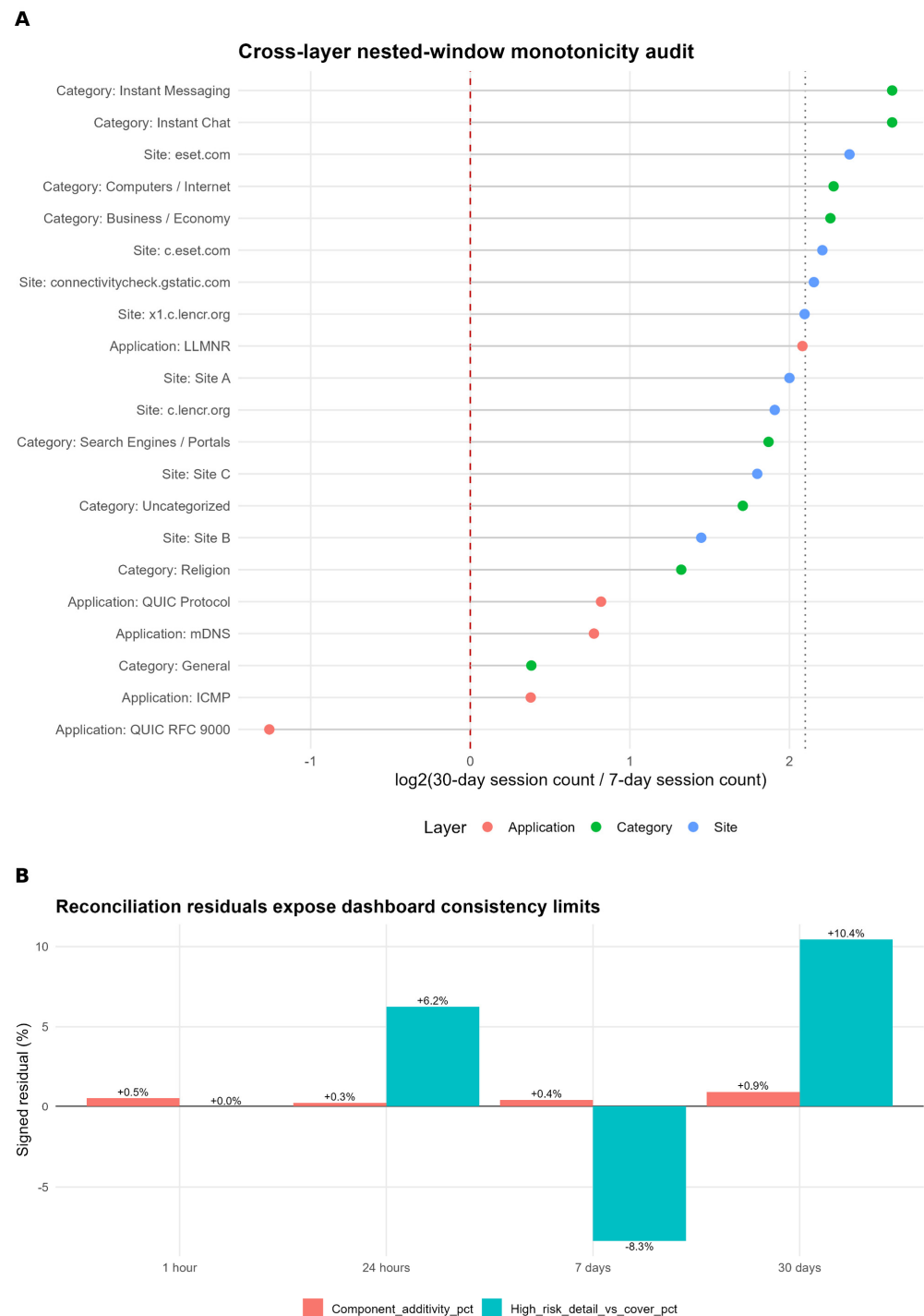


Figure 6. Cross-horizon and within-report consistency audit. **(A)** $\log_2(30\text{-day}/7\text{-day})$ for session metrics shared by the nested 7-day and 30-day reports; dashed line, strict monotonicity boundary; dotted line, proportional-exposure reference. **(B)** Signed reconciliation residuals for component additivity (salmon) and detailed-versus-cover high-risk counts (teal).

Table 4. Reconciliation Audit: These inconsistencies do not make the reports unusable; they define the confidence boundary and the need for raw evidence.

Check	Observed Discrepancy	Security Implication	Analytic Treatment
Component additivity	Received + sent differs from cover total by 0.5–0.9% in several reports	Rates and shares can shift if mixed denominators are used	Use cover total for rate; component sum for direction; disclose rounding

Hourly bin coverage	Detailed bins total 402 MB versus 421 MB on cover	Edge activity is omitted or governed by different inclusion rules	Treat binned profile as 95.5% coverage, not a full ledger
High-risk KPI	Cover/detailed counts: 16/17, 108/99, 833/920	Metric semantics or refresh timing differ within a report	Keep metrics separate; do not aggregate them as one variable
Cross-horizon QUIC sessions	Weekly QUIC (RFC 9000) = 170 K; 30-day value = 71 K	Violates monotonicity under stable taxonomy and aligned containment	Flag taxonomy, ranking, cache, or report-generation drift; require raw logs
Infected-device summary	Cover reports zero infected devices, while weekly and 30-day detail lists two and four hosts	Status KPI and historical incident list use different definitions or states	Describe as unresolved semantic conflict; do not infer active infection

The nested-window monotonicity audit provides the strongest consistency test (Figure 6A). For a cumulative session count with stable taxonomy and inclusion rules, a 30-day window containing the final seven days cannot report fewer sessions than the 7-day window. Most shared site and category counts satisfied this minimal condition even when their growth was far from proportional to the elapsed time. QUIC (RFC 9000) did not: the displayed count fell from 170 K sessions in seven days to 71 K in the containing 30-day report, giving $\log_2(71/170) \approx -1.26$ and crossing the strict monotonicity boundary. In Figure 6A, values below zero are inconsistent with simple containment, whereas the proportional-exposure reference at $\log_2(30/7)$ is shown only for orientation. The generic “Quic Protocol”, mDNS, ICMP, and LLMNR rows remained non-negative. The reversal therefore cannot be explained by ordinary temporal sparsity; it requires a change in taxonomy, ranking, caching, deduplication, report-generation logic, or a reporting defect. The available exports cannot distinguish these candidate mechanisms, so the discrepancy is reported as unresolved rather than explained. The correct analytic response is to preserve the discrepancy and request raw logs plus the vendor metric definition rather than choose the more convenient value. Figure 6B shows a second class of disagreement. Component additivity residuals were small (+0.3% to +0.9%, consistent with display rounding), whereas high-risk detail-versus-cover residuals reached +6.2% at 24 h, −8.3% at seven days, and +10.4% at 30 days. These residuals are not “error rates”; they quantify the degree to which dashboard fields that appear comparable cannot safely be collapsed into one variable.

4.7. Alignment-Controlled and Resolution-Harmonised Sensitivity Analyses

Because the one-hour and daily windows are offset from the 7-day/30-day pair (Section 3.2), differences involving them combine a reporting-horizon effect with a time-of-observation effect. Only one report of each duration was available, so we could not construct repeated rolling windows; instead, three analyses separate these components using the exactly nested pair and matched within-report distributions. First, splitting the 30-day report into the final week and the preceding 23 days yields disjoint, contiguous, and aligned windows. Their duration-normalised traffic rates were 3.51 and 3.54 GB/day (ratio 0.99), and their received shares were 69.2% and 71.0%. At the weekly scale, workload volume and direction were therefore effectively stationary, and the large 7-day/30-day differences in application concentration and security-event counts cannot be explained by a change in total workload. Security-control intensity was not stationary, however: the final week recorded 2234 IPS incidents per day versus 5235 per day in the preceding 23 days (ratio 0.43), and 0.57 versus 1.04 Anti-Virus events per day. Event emergence across horizons is therefore a joint product of longer exposure and genuine temporal variation in control activity.

Second, we compared the short windows with matched distributions of equal-length intervals rather than a single long-run mean. The one-hour value (0.421 GB) was 1.18 standard deviations above the mean of the 24 one-hour bins in the daily report (mean 0.161 GB, SD 0.220 GB). The daily value (3.86 GB) was 0.15 standard deviations above the mean of the 30 daily bins in the 30-day report (mean 3.53 GB, SD 2.19 GB) and approximately 1.19 times their median of about 3.25 GB. The 2.86-fold duration-normalised contrast for the final hour therefore reflects sampling one interval from a strongly diurnal and bursty process: it is elevated relative to the 30-day mean but lies within the ordinary hour-to-hour variation in the preceding day, and the final day was typical. These comparisons use report-native bins from adjacent rather than identical periods and are reference distributions, not significance tests.

Third, we assessed the sensitivity of the burst metrics to bin width using the resolution-harmonised coefficient of variation (Section 3.4). Rescaled to a common 24 h resolution under an independence assumption, CV^* was 0.05 for the one-hour report, 0.28 for the daily report, 0.56 for the weekly report, and 0.62 for the 30-day report. The ordering therefore reverses relative to the native-resolution metrics, in which the daily report appeared most bursty. Across bin widths of 1 to 24 h, CV decreased approximately to $\Delta^{-0.24}$, about half the $\Delta^{-0.5}$ rate expected for independent bins. This slope should not be read as evidence of temporal dependence. The three points come from different windows, so the slope also reflects differences in the workload each window observed: the daily window is dominated by a single diurnal cycle, whereas the 30-day window mixes weekday, weekend, and maintenance regimes, and such non-stationarity can by itself produce the slow decay of variability with aggregation that is otherwise attributed to long-range dependence [37]. The result is compatible with, but does not test, the long-range-dependent variability documented for local-area and wide-area traffic [38,39]; a dependence claim would require a variance–time analysis across several aggregation levels within a single window [38]. The exponent is therefore not reported as a Hurst estimate. Cross-horizon burstiness rankings in these data are therefore not resolution-invariant, and CV^* is interpreted only as a diagnostic of resolution sensitivity; native-resolution CV and peak-to-median values are therefore reported only as within-window descriptors, and full harmonisation requires re-binning time-stamped records to a common resolution.

We then tested the exposure-adjusted localisation metric against non-uniform baselines. The 7-day window contains exactly one complete weekly cycle, and the 30-day window contains four complete cycles plus a 48 h Wednesday–Friday remainder. For any exposure profile that repeats weekly, the expected final-week share therefore lies between 20% and 25%, whatever its weekday, weekend, or business-hour pattern; weekday-only, business-hours-only (Monday–Friday, 08:00–18:00), and uniform profiles gave 22.7%, 22.7%, and 23.3%, respectively. Across the full 20–25% band, localisation remained indistinguishable from proportional for total traffic ($L = -0.11$ to $+0.21$), negative for IPS incidents (-1.12 to -0.80), detailed high-risk sessions (-1.22 to -0.89), TeamViewer (-0.70 to -0.38), and WireGuard (-1.72 to -1.40), and positive for uncategorised traffic ($+0.32$ to $+0.64$) and TunnelBear ($+2.00$ to $+2.32$). Small counts impose a second limit. Under an idealised independent-event model, exact (Clopper–Pearson [40]) 95% binomial intervals for the final-week share were 4.0–32.7% for Anti-Virus events (4/28), 6.8–93.2% for hosts (2/4), 39.8–100% for TunnelBear (4/4), 2.5–16.8% for WireGuard (5/66), 12.6–18.6% for TeamViewer (90/583), 8.8–12.9% for detailed high-risk sessions (99/920), and 11.3–11.7% for IPS incidents (15,641/136,057). The Anti-Virus and host intervals include the 20–25% band and are not distinguishable from weekly-periodic exposure even under independence. Security events are, however, clustered in time, as the zero IPS count in the 23 h daily overlap shows (next paragraph), so independent-event intervals understate uncertainty. We therefore added a clustering-aware sensitivity analysis based on the design-effect

adjustment for clustered binary data [41]: both the final-week count and the 30-day total are divided by a design effect d before the exact interval is computed. Here d can be read as the mean number of events that share a common cause, and therefore a common temporal location; $d = 1$ recovers the independent-event case. For each indicator, we report the smallest d at which the interval reaches the 20–25% band. TunnelBear and WireGuard reached the band at $d \approx 1.6$ and TeamViewer at $d \approx 2.0$, so their apparent localisation is not robust to even mild clustering. Detailed high-risk sessions remained outside the band up to $d \approx 12$, equivalent to about 75 independent episodes over 30 days, and IPS incidents up to $d \approx 1500$, equivalent to about 90 episodes; the latter is not implausible, because a single repeated protection, TCP Segment Limit Enforcement, dominated the IPS tables. If the number of independent episodes were limited to one per day ($d = n/30$), none of the indicators with more than 30 events would remain distinguishable. The signs of the localisation ratios are therefore reported as descriptive; only the final-week deficits in IPS incidents and detailed high-risk sessions survive moderate clustering, and even these depend on a cluster structure that only time-stamped event records can establish. Localisation is accordingly interpreted as deviation from weekly-periodic exposure, not as evidence of recent abnormality; monthly maintenance or patch cycles are not covered by this bound.

The same reasoning explains why a short quiet window is weak evidence of security. The daily window shares 23 h with the weekly window. Had the weekly IPS incidents occurred at a constant rate, approximately 2140 would be expected in that overlap; the reported zero therefore implies either strong temporal clustering of IPS activity, as is typical of non-Poisson network processes [39], or different inclusion rules in the short-window cover summaries. For rare events, a zero is uninformative even without clustering: at the weekly Anti-Virus rate, the probability of observing no event in 23 h is 0.58, and at the 30-day Anti-Bot rate the probability of observing no event in 24 h is 0.94.

Table 5 consolidates these checks by stating, for each principal finding, whether it is attributable to temporal aggregation, to time of observation, or to vendor reporting logic. One additional inconsistency emerged from the containment bound. QUIC (RFC 9000) was the leading weekly application at approximately 30% of 24.6 GB, or about 7.4 GB, yet the leading 30-day application (Windows 11 update) accounted for only about 5.2 GB (4.9% of 106 GB). Under a stable taxonomy, QUIC's 30-day volume could not be lower than its weekly volume and would itself exceed 7.0% of 30-day traffic. The top-1 bound is therefore violated, mirroring the session-count reversal in Figure 6A, whereas the top-five bound ($\geq 14.3\%$ versus 17.5% observed) and the top-ten bound ($\geq 18.3\%$ versus 25%) are satisfied. Part of the apparent collapse in leading-application dominance between the 7-day and 30-day views therefore reflects reporting logic rather than aggregation. Of the six material inconsistencies now identified (the five in Table 4 and the top-1 bound), two, the QUIC session reversal and the QUIC volume bound, are detectable only through cross-horizon containment checks and would pass single-report completeness, accuracy, and additivity checks. This is the measurable added value of the scale-aware checks relative to single-artefact data-quality assessment [31–33].

Table 5. Attribution of principal findings. “Aggregation” denotes an effect of reporting horizon under stable definitions; “time of observation” an effect of when a window was observed; and “reporting logic” an effect of vendor metric definitions, taxonomy, or report generation.

Finding	Alignment-Controlled Evidence	Primary Attribution	Evidential Status
Final-hour 2.86-fold rate contrast	Hourly window disjoint from 7-day/30-day pair; +1.18 SD within preceding day's one-hour bins	Time of observation (one interval of a bursty diurnal process)	Observation

Cross-horizon burstiness ordering	Ordering reverses after resolution harmonisation (CV* 0.05–0.62)	Bin width, not separable from window-specific workload (resolution-sensitivity diagnostic)	Within-window descriptor only
Decline in top-five and top-ten concentration	Containment bounds satisfied; aligned traffic rates equal (ratio 0.99)	Consistent with aggregation (workload mixing); not uniquely identified	Observation
Decline in top-1 share (30% to 4.9%)	Containment bound violated (≈ 7.4 GB weekly QUIC > ≈ 5.2 GB 30-day leader)	Reporting logic	Data-quality flag
Emergence of IPS and Anti-Virus events	Final-week IPS rate $0.43 \times$ preceding 23 days; zero in 23 h daily overlap	Aggregation plus time of observation (clustered events)	Level A (control activity)
Weekly > 30-day QUIC sessions	Violates nested monotonicity	Reporting logic (mechanism unresolved)	Data-quality flag
High-risk cover/detail residuals; infected-device conflict	Within-report; independent of horizon	Reporting logic (metric semantics)	Data-quality flag
Final-week localisation (uncategorised traffic, TunnelBear)	Sign stable for any weekly-periodic baseline (20–25%); TunnelBear not robust to mild clustering ($d \approx 1.6$)	Time of observation (variation within 30 days)	Level B indicator

5. Discussion

5.1. Reporting Horizon Is a Measurement Operator, Not a Neutral Display Setting

RQ1 asked how traffic characteristics change across horizons. The central finding is not simply that longer reports contain more traffic. The reporting horizon changes the workload's apparent structure. The final hour operated at 2.86 times the 30-day normalised rate, yet the 24 h report exhibited the strongest within-window burstiness. At the same time, application concentration collapsed as the horizon expanded: the top application fell from 70% of hourly traffic to 4.9% of 30-day traffic, and the cumulative top-10 share fell from 96% to 25%. These are different manifestations of the same measurement effect. The alignment-controlled analysis (Table 5) qualifies this interpretation: the final-hour contrast is largely a time-of-observation effect, the burstiness ordering is sensitive to bin width and cannot be separated from window-specific workload, and part of the top-application collapse reflects reporting logic, whereas the decline in top-five and top-ten concentration remains consistent with aggregation. Short windows magnify the applications and bursts active “now”; long windows mix regimes and expose diversity.

The Jaccard analysis adds an important qualification. Adjacent operational horizons can preserve application membership while reweighting its importance: the 24 h and 7-day top-10 sets had $J = 0.82$, yet their concentration profiles differed visibly. By contrast, the one-hour and 30-day sets shared little membership ($J = 0.25$). A dashboard can therefore change the risk narrative through two mechanisms: rank-weight distortion, in which the same applications appear with different dominance, and membership churn, in which different applications enter the analyst's top-k view. Treating a single horizon as the network's “application profile” obscures both mechanisms.

These results are consistent with multi-scale anomaly research [11–13] but add an operational measurement perspective. Multi-scale models assume that the underlying observations can be compared or fused. Vendor reports first require evidence governance:

time alignment, definition control, top-k awareness, and consistency testing. More scales do not automatically produce more truth if the scales encode different semantics.

5.2. Evidence Emergence and Temporal Localisation Reshape the Security Narrative

RQ2 concerned event visibility and prioritisation. The security-evidence heatmap makes the asymmetry clear: zero Anti-Bot, Anti-Virus, and IPS counts in the fine windows coexist with 15,641 weekly and 136,057 30-day IPS incidents. This is not a contradiction; the events may simply fall outside the shorter interval. Section 4.7 shows that this emergence combines longer exposure with genuinely lower control activity in the final week and with strong temporal clustering of IPS events. The operational problem is interpretive. A fine-window green status can answer “what was recorded in this interval,” but it cannot answer “is the environment secure?” without a longer event history, unresolved investigations, control health, and corroborating telemetry.

The localisation analysis goes beyond that visibility result by showing where activity sat within the 30-day history. Network volume itself accumulated almost exactly in proportion to elapsed time, while IPS, Anti-Virus, detailed high-risk, TeamViewer, and WireGuard activity were relatively sparse in the final week. Uncategorized traffic and sessions were comparatively concentrated in the final week, and all displayed TunnelBear activity fell within it. This distinction matters for triage: a high monthly count can describe persistent historical burden, whereas a positive localisation ratio identifies a deviation from weekly-periodic exposure that warrants triage, not evidence of recent abnormality. Neither measure alone establishes maliciousness, and sparse counts require particular caution. Because the clustering-aware analysis in Section 4.7 shows that most localisation contrasts are not robust to even mild clustering of events, the ratios are used to order triage rather than to claim that activity was distinguishably concentrated in time.

5.3. From Dashboard Telemetry to Confidence-Aware Cyber-Risk Intelligence

RQ3 asked which conditions must be satisfied before dashboard observations can support cyber-risk decisions. Figure 6 shows why reconciliation belongs inside the analytical workflow rather than in a data-cleaning appendix. The QUIC RFC 9000 reversal violates a minimal property of nested cumulative counts, while the high-risk cover/detail residuals show that visually adjacent dashboard fields may use different semantics or refresh states. A mature workflow should therefore treat consistency failures as first-class evidence about confidence. The following sequence is a proposed best-practice workflow informed by these findings and the CTI literature [25–30]; unlike the results in Section 4, this study has not tested its operational effectiveness.

1. Preserve provenance. Retain raw events alongside report exports; record appliance version, timezone, collection interval, filter state, policy version, and report-generation time.
2. Align time. Convert timestamps to one timezone, document partial bins, and compare only genuinely nested intervals. Distinguish event time from report-generation time.
3. Normalise exposure. Express traffic and event counts per unit time and, when available, per active asset, user, or session. Keep raw counts visible so normalisation does not conceal operational load.
4. Reconcile semantics. Test component additivity, nested-window monotonicity, top-k censoring, taxonomy stability, and agreement between cover KPIs and detailed rows. Distinguish harmless display rounding from contradictions that change interpretation; route unresolved failures to data engineering or the vendor before aggregation or modelling.

5. Establish baselines. Compare current values with matched day-of-week, hour-of-day, workload, and change-window baselines. Use robust statistics and maintain separate baselines for known operating regimes.
6. Corroborate indicators. Link gateway observations to endpoint, identity, DNS, vulnerability, email, cloud control plane, and change-management evidence. A burst plus a malware alert plus an unexpected process and destination is stronger than any one signal.
7. Grade confidence. Label each analytic claim with one of the four ordered evidence states (observation, indicator, control-confirmed event, or analyst-confirmed incident) using the transition rules in Section 3.5, and attach candidate explanations separately as hypotheses. Record disconfirming evidence and benign alternatives.
8. Map to risk and action. Connect the evidence to an asset, business service, threat scenario, control objective, owner, response threshold, and expiry time. Intelligence that does not change a decision is an information product, not yet a risk product.

Table 6 shows how the supplied signals can be operationalised without overclaiming.

Table 6. Evidence-to-Action Matrix: Each action is conditioned on context and corroboration rather than the dashboard label alone.

Signal	Plausible Benign Explanation	Plausible Threat Hypothesis	Evidence Needed Next	Proportionate Action
Short-window traffic burst	Update, backup, synchronisation, user demand	Flooding, payload delivery, bulk extraction	Flow records, endpoints, destinations, service health, change calendar	Triage and correlate; rate-limit only if service/policy evidence supports it
High QUIC volume/sessions	Normal web and application transport	Policy bypass or concealed malicious traffic	SNI/metadata where lawful, endpoint process, destination reputation, baseline	Validate policy and application identity; do not block solely by protocol
High LLMNR/mDNS sessions	Local discovery and name resolution	Spoofing, poisoning, reconnaissance	Source hosts, failed lookups, responder identity, endpoint telemetry	Reduce unnecessary protocols; investigate anomalous responders
Uncategorized traffic	New or unsupported application signature	Obfuscated protocol or command-and-control	Packet/flow metadata, destination ownership, application update status	Improve classification; isolate only after corroboration
TeamViewer/VPN/BitTorrent	Approved support, privacy, or transfer use	Unauthorised remote control, tunnelling, exfiltration	Software inventory, owner approval, authentication and endpoint logs	Enforce allow-list and least privilege; investigate unauthorised use
IPS/Anti-Virus event	Blocked exploit attempt or false positive	Exploitation or malware execution	Raw alert, signature, action, host state, repeat pattern	Validate disposition; escalate confirmed compromise through IR process

5.4. Recommendations for Security Architecture and Governance

The findings suggest three recommendations for security operations; they are derived from, rather than tested by, the present data. First, dashboards should expose multiple horizons concurrently rather than force analysts to choose one. A useful design pairs a fine-grained burst panel with normalised longer-run context, rolling security-event history, recent-localisation indicators, and an unresolved-investigation queue. Second, report generators need machine-readable exports and explicit metric contracts. Every field should specify its scope, unit, rounding rule, deduplication logic, update latency, taxonomy version, and whether top-k selection occurs before or after filtering. Consistency tests such as component additivity, top-k state tracking, and nested-window monotonicity can

then run automatically. A reversal such as 170 K weekly versus 71 K 30-day QUIC RFC 9000 sessions should generate a data-quality alert before the value reaches an executive risk report.

Third, governance must connect monitoring to business risk. NIST continuous-monitoring guidance emphasises visibility into assets, threats, vulnerabilities, and control effectiveness [1], while CSF 2.0 places detection and response within a broader governance cycle [2]. The supplied reports are strongest as Detect Evidence and as prompts for Protect and Respond activity. Their evidential value changes materially when enriched with asset criticality, service dependencies, approved-use records, change windows, and control outcomes. The proposed workflow makes that transformation explicit by attaching provenance, temporal context, confidence, competing explanations, and an action threshold to each signal.

5.5. Relation to Prior Work and Extension to Distributed Wireless Environments

The study supports the premise that temporal context and multiple views improve anomaly analysis [7,11–13], but it differs from classifier-centred work by examining the measurement properties of a low-resolution practitioner artefact before modelling. No predictive accuracy is claimed because there are no labels or raw samples. The contribution is instead a set of reproducible diagnostics for whether aggregated operational evidence is stable enough to interpret: rate normalisation, within-window intensity, cumulative concentration, application-set similarity, exposure localisation, monotonicity, and reconciliation residuals. The results also reinforce warnings about closed-world assumptions and evaluation bias [14,15,17]. A traffic peak, high QUIC volume, LLMNR activity, or use of a risk-labelled application can support investigation, but the same observation can have legitimate explanations. Without ground truth, the scientifically defensible output is a bounded set of hypotheses with explicit confidence, not an attack label.

Finally, the analysis extends the CTI practice literature from external threat reports to internally generated operational evidence. Multi-granular fusion methods can integrate disparate threat descriptions [28,29], and privacy-preserving approaches can support cross-organisational correlation [30]. This study identifies a prerequisite for such fusion: local telemetry must first satisfy provenance and consistency checks. Reconciliation is therefore not merely a data-engineering concern; it is part of the evidential chain that determines whether a technical observation can become trustworthy intelligence.

The methodology is not specific to a single wired edge gateway. Emerging low-altitude and beyond-5G wireless networks distribute reception across multiple access points whose cooperation and duplex modes change dynamically; in flexible-duplex cell-free architectures, for example, distributed access points switch between receiving uplink signals and transmitting cooperative artificial noise to secure uplink communications [42]. In such settings, each node produces its own telemetry with its own clock, reporting interval, and operating mode. The alignment and reconciliation steps extend directly: node reports must be aligned with a common time base; nested-window and cross-node containment constraints, such as the sum of per-node counts over a common interval versus a network-level aggregate, must be tested; and mode changes must be retained as provenance so that a change in cooperation mode is not mistaken for a change in threat activity. Multifunctional wireless systems also broaden the available evidence. Intelligent-reflection technologies that integrate communication, sensing, and security [43] can provide channel-state, sensing, and secrecy-performance observations alongside conventional network counters. The evidence levels in Section 3.5 can accommodate such heterogeneous cross-layer observations, treating a sensing anomaly or a drop in secrecy performance as a Level B indicator until corroborated; however, each source requires its own metric contract and time-

alignment rules before fusion. Empirical evaluation of these extensions is beyond the scope of the present single-gateway study.

5.6. Limitations and Threats to Validity

This study has important limitations. It covers one gateway and, at most, 30 days, so it cannot estimate prevalence or generalise event rates to other enterprises; its findings are case-study findings rather than general properties of security dashboards. Only one report of each duration was available, so horizon effects could not be replicated across multiple rolling windows; the aligned split and matched within-report distributions in Section 4.7 partially substitute. The one-hour and 24 h windows are offset by up to 111 min relative to the 7-day/30-day pair, and the one-hour window does not overlap with that pair; consequently, strict nested-window tests apply only to the aligned 7-day and 30-day reports. The PDFs expose rounded aggregates and top-k lists rather than complete inventories of packets, flows, alerts, or applications. Jaccard similarity therefore describes the stability of the displayed top-10 set, not the full latent application population, and “not listed” remains censored rather than zero.

The reports provide no labelled ground truth, attack diary, incident ticket, business calendar, network topology, asset inventory, user context, or pre-period baseline. Consequently, these reports alone cannot attribute a traffic burst, protocol, site, malware row, or high-risk application to an attack. Conversely, zero fine-window events cannot rule out undetected activity. Control counts may represent blocked attempts, repeated signatures, sessions, deduplicated incidents, or report-specific units.

The analysis relied on Check Point appliance-rendered reports. We double-checked values against the source pages, but display rounding creates unavoidable non-additivity, and transcription error cannot be eliminated entirely. Report-native bins differ, so burst statistics are used only as within-window descriptors; Section 4.7 shows that their cross-horizon ordering is not resolution-invariant, and the resolution-harmonised CV* is used only as a diagnostic of resolution sensitivity because the windows differ in the workload they observe. Because the reports are vendor-generated summaries rather than packet, flow, or event records, we could not attribute the discrepancies in Table 5 to a specific reporting mechanism. Exposure-adjusted localisation assumes comparable metric semantics across the exactly nested 7-day and 30-day views; the monotonicity audit demonstrates why that assumption must be tested rather than taken for granted. Its uncertainty also depends on how events cluster in time; the design-effect sensitivity analysis in Section 4.7 bounds this dependence but cannot estimate it without time-stamped event records. This study does not evaluate detection accuracy, the causal effects of an intervention, or the superiority of one deployment architecture over another.

The confidence framework has not been independently validated. Its operational value should be tested in three steps: (i) an inter-analyst reliability study in which at least three analysts independently assign the four ordered evidence states to the same set of dashboard findings, summarised with Krippendorff’s α for ordinal data, and independently code the attached hypotheses, summarised with α for nominal data, using $\alpha \geq 0.80$ as the conventional threshold for reliable coding and $\alpha \geq 0.667$ as the minimum for tentative conclusions [44], reporting agreement separately for each transition rule in Section 3.5 so that inconsistently applied rules can be identified; (ii) retrospective validation against analyst-adjudicated incident tickets, testing whether higher evidence levels are associated with higher confirmed-incident rates; and (iii) controlled benign-workload and authorised attack scenarios. Until such evidence is available, the framework should be regarded as a structured analytic convention rather than a validated cyber-risk instrument.

External validity is analytical rather than statistical: other security teams can apply the measurement and reconciliation workflow to their own dashboard exports, but they should not transfer the observed traffic rates, application mix, and event counts. A stronger follow-up study should obtain raw event logs over at least 90 days, align all sources with UTC, preserve policy/signature/taxonomy versions, define asset-aware baselines, and have analysts adjudicate a stratified sample of alerts. Controlled benign workload changes and authorised attack simulations could then estimate precision, recall, time-to-detection, and time-to-triage while directly testing whether the dashboard-level diagnostics predict evidence quality.

6. Conclusions

In this single-gateway case study, four views of the same gateway produced materially different security narratives. Alignment-controlled analysis showed three distinct sources of these differences: temporal aggregation, the time at which each window was observed, and vendor reporting logic. The final hour operated at 2.86 times the 30-day normalised traffic rate, although within the ordinary hour-to-hour variation in the preceding day; burstiness rankings were sensitive to bin width; and top-10 application concentration fell from 96% to 25% as the horizon expanded. Top-10 membership also depended on horizon (Jaccard 0.25–0.82). Fine-window silence in Anti-Bot, Anti-Virus, and IPS controls coexisted with substantially longer-window activity, while exposure-adjusted localisation described where activity sat within the 30-day history, although most localisation contrasts were not robust to the temporal clustering of events. At the same time, the dashboard failed several consistency checks, most notably the 170 K weekly versus 71 K 30-day QUIC RFC 9000 reversal and high-risk cover/detail residuals approaching 10%.

The central implication is that network telemetry does not become cyber-risk intelligence through aggregation alone. Reporting horizon acts as a measurement operator: it changes burst visibility, application dominance, top-k membership, security-evidence emergence, and the apparent recency of activity. Cyber-risk decisions should therefore be made only after time alignment, exposure normalisation, semantic reconciliation, corroboration, competing-explanation analysis, and explicit confidence grading. Short windows are valuable for triage, long windows are essential for context and persistence, and nested-window consistency tests provide a practical safeguard against converting dashboard artefacts into false certainty. The proposed workflow offers a candidate route from vendor reporting to more scientifically defensible and decision-relevant cyber-risk intelligence; its operational value now requires independent validation across multiple gateways, longer periods, and adjudicated incidents.

Author Contributions: Conceptualization, M.R.H.; Methodology, M.R.H. and N.I.S.; Software, M.R.H. and N.I.S.; Validation, M.R.H., N.I.S. and N.H.S.A.; Formal analysis, M.R.H.; Investigation, M.R.H.; Resources, N.I.S., N.H.S.A. and M.R.H.; Data curation, M.R.H.; Writing—original draft, M.R.H.; Writing—review and editing, N.I.S. and N.H.S.A.; Visualisation, M.R.H.; Supervision, N.I.S., N.H.S.A. and M.R.H.; Project administration, N.I.S. and N.H.S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Data Availability Statement: The study analyses proprietary operational gateway reports containing security-sensitive infrastructure identifiers. The raw reports are therefore not publicly available. De-identified aggregate values required to reproduce the reported analyses are provided in the article and can be supplied with the accompanying R script, subject to organisational approval.

Acknowledgments: The authors acknowledge the organisation BICC Cyber Security & IT Solutions, which supplied the operational gateway reports; Tahlina Ahmed, the director of the company, who authorised and consented us to use the data; and the technical staff, Masudur Rahman, who supported de-identification and verification.

Conflicts of Interest: The authors declare no known competing financial interests or personal relationships that could have influenced the work reported. Check Point Software Technologies Ltd. supplied no funding and had no role in study design, analysis, interpretation, manuscript preparation, or the decision to submit.

References

1. *NIST SP 800-137; Information Security Continuous Monitoring (ISCM) for Federal Information Systems and Organizations*. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2011. <https://doi.org/10.6028/NIST.SP.800-137>.
2. *NIST CSWP 29; The NIST Cybersecurity Framework (CSF) 2.0*. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2024. <https://doi.org/10.6028/NIST.CSWP.29>.
3. *NIST SP 800-61 Rev. 3; Incident Response Recommendations and Considerations for Cybersecurity Risk Management: A CSF 2.0 Community Profile*. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2025. <https://doi.org/10.6028/NIST.SP.800-61r3>.
4. *NIST SP 800-94; Guide to Intrusion Detection and Prevention Systems (IDPS)*. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2007. <https://doi.org/10.6028/NIST.SP.800-94>.
5. *NIST SP 800-92; Guide to Computer Security Log Management*. National Institute of Standards and Technology: Gaithersburg, MD, USA, 2006. <https://doi.org/10.6028/NIST.SP.800-92>.
6. Paxson, V. Bro: A system for detecting network intruders in real-time. *Comput. Netw.* **1999**, *31*, 2435–2463. [https://doi.org/10.1016/S1389-1286\(99\)00112-7](https://doi.org/10.1016/S1389-1286(99)00112-7).
7. Lakhina, A.; Crovella, M.; Diot, C. Diagnosing network-wide traffic anomalies. In Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Portland, OR, USA, 30 August–3 September 2004; pp. 219–230. <https://doi.org/10.1145/1015467.1015492>.
8. García-Teodoro, P.; Díaz-Verdejo, J.; Maciá-Fernández, G.; Vázquez, E. Anomaly-based network intrusion detection: Techniques, systems and challenges. *Comput. Secur.* **2009**, *28*, 18–28. <https://doi.org/10.1016/j.cose.2008.08.003>.
9. Mirsky, Y.; Doitshman, T.; Elovici, Y.; Shabtai, A. Kitsune: An ensemble of autoencoders for online network intrusion detection. In Proceedings of the 2018 Network and Distributed System Security Symposium, San Diego, CA, USA, 18–21 February 2018. <https://doi.org/10.14722/ndss.2018.23204>.
10. Shone, N.; Ngoc, T.N.; Phai, V.D.; Shi, Q. A deep learning approach to network intrusion detection. *IEEE Trans. Emerg. Top. Comput. Intell.* **2018**, *2*, 41–50. <https://doi.org/10.1109/TETCI.2017.2772792>.
11. Duan, X.; Fu, Y.; Wang, K. Network traffic anomaly detection method based on multi-scale residual classifier. *Comput. Commun.* **2023**, *198*, 206–216. <https://doi.org/10.1016/j.comcom.2022.10.024>.
12. Cohen, S.; Goldshlager, N.; Shapira, B.; Rokach, L. TTANAD: Test-time augmentation for network anomaly detection. *Entropy* **2023**, *25*, 820. <https://doi.org/10.3390/e25050820>.
13. Liu, T.; Fu, Y.; Wang, K.; Duan, X.; Wu, Q. A multiscale approach for network intrusion detection based on variance-covariance subspace distance and EQL v2. *Comput. Secur.* **2025**, *148*, 104173. <https://doi.org/10.1016/j.cose.2024.104173>.
14. Sommer, R.; Paxson, V. Outside the closed world: On using machine learning for network intrusion detection. In Proceedings of the 2010 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 16–19 May 2010; pp. 305–316. <https://doi.org/10.1109/SP.2010.25>.
15. Arp, D.; Quiring, E.; Pendlebury, F.; Warnecke, A.; Pierazzi, F.; Wressnegger, C.; Cavallaro, L.; Rieck, K. Dos and don'ts of machine learning in computer security. In *31st USENIX Security Symposium (USENIX Security 22)*; USENIX Association: Berkeley, CA, USA, 2022; pp. 3971–3988. Available online: <https://www.usenix.org/conference/usenixsecurity22/presentation/arp> (accessed on 27 August 2026).
16. Buczak, A.L.; Guven, E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 1153–1176. <https://doi.org/10.1109/COMST.2015.2494502>.
17. Pendlebury, F.; Pierazzi, F.; Jordaney, R.; Kinder, J.; Cavallaro, L. TESSERACT: Eliminating experimental bias in malware classification across space and time. In *28th USENIX Security Symposium (USENIX Security 19)*; USENIX Association: Berkeley, CA,

- USA, 2019; pp. 729–746. Available online: <https://www.usenix.org/conference/usenixsecurity19/presentation/pendlebury> (accessed on 27 August 2026).
18. Sharafaldin, I.; Lashkari, A.H.; Ghorbani, A.A. Toward generating a new intrusion detection dataset and intrusion traffic characterization. In Proceedings of the 4th International Conference on Information Systems Security and Privacy, Funchal, Portugal, 22–24 January 2018; pp. 108–116. <https://doi.org/10.5220/0006639801080116>.
 19. Ring, M.; Wunderlich, S.; Scheuring, D.; Landes, D.; Hotho, A. A survey of network-based intrusion detection data sets. *Comput. Secur.* **2019**, *86*, 147–167. <https://doi.org/10.1016/j.cose.2019.06.005>.
 20. Khraisat, A.; Gondal, I.; Vamplew, P.; Kamruzzaman, J. Survey of intrusion detection systems: Techniques, datasets and challenges. *Cybersecurity* **2019**, *2*, 20. <https://doi.org/10.1186/s42400-019-0038-7>.
 21. Koroniotis, N.; Moustafa, N.; Sitnikova, E.; Turnbull, B. Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset. *Future Gener. Comput. Syst.* **2019**, *100*, 779–796. <https://doi.org/10.1016/j.future.2019.05.041>.
 22. Bridges, R.A.; Glass-Vanderlan, T.R.; Iannacone, M.D.; Vincent, M.S.; Chen, Q. A survey of intrusion detection systems leveraging host data. *ACM Comput. Surv.* **2019**, *52*, 128. <https://doi.org/10.1145/3344382>.
 23. Cheng, T.; Adepeju, M. Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *PLoS ONE* **2014**, *9*, e100465. <https://doi.org/10.1371/journal.pone.0100465>.
 24. Landauer, M.; Onder, S.; Skopik, F.; Wurzenberger, M. Deep learning for anomaly detection in log data: A survey. *Mach. Learn. Appl.* **2023**, *12*, 100470. <https://doi.org/10.1016/j.mlwa.2023.100470>.
 25. Tounsi, W.; Rais, H. A survey on technical threat intelligence in the age of sophisticated cyber attacks. *Comput. Secur.* **2018**, *72*, 212–233. <https://doi.org/10.1016/j.cose.2017.09.001>.
 26. Ainslie, S.; Thompson, D.; Maynard, S.; Ahmad, A. Cyber-threat intelligence for security decision-making: A review and research agenda for practice. *Comput. Secur.* **2023**, *132*, 103352. <https://doi.org/10.1016/j.cose.2023.103352>.
 27. Shin, B.; Lowry, P.B. A review and theoretical explanation of the ‘Cyberthreat-Intelligence (CTI) capability’ that needs to be fostered in information security practitioners and how this can be accomplished. *Comput. Secur.* **2020**, *92*, 101761. <https://doi.org/10.1016/j.cose.2020.101761>.
 28. Ma, C.; Jiang, Z.; Zhang, K.; Ling, Z.; Jiang, J.; You, Y.; Yang, P.; Feng, H. TIMFuser: A multi-granular fusion framework for cyber threat intelligence. *Comput. Secur.* **2025**, *148*, 104141. <https://doi.org/10.1016/j.cose.2024.104141>.
 29. Wang, G.; Liu, P.; Huang, J.; Bin, H.; Wang, X.; Zhu, H. KnowCTI: Knowledge-based cyber threat intelligence entity and relation extraction. *Comput. Secur.* **2024**, *141*, 103824. <https://doi.org/10.1016/j.cose.2024.103824>.
 30. Preuveneers, D.; Joosen, W. Privacy-preserving correlation of cross-organizational cyber threat intelligence with private graph intersections. *Comput. Secur.* **2023**, *135*, 103505. <https://doi.org/10.1016/j.cose.2023.103505>.
 31. Wang, R.Y.; Strong, D.M. Beyond accuracy: What data quality means to data consumers. *J. Manag. Inf. Syst.* **1996**, *12*, 5–33. <https://doi.org/10.1080/07421222.1996.11518099>.
 32. Schlette, D.; Böhm, F.; Caselli, M.; Pernul, G. Measuring and visualizing cyber threat intelligence quality. *Int. J. Inf. Secur.* **2021**, *20*, 21–38. <https://doi.org/10.1007/s10207-020-00490-y>.
 33. Zibak, A.; Sauerwein, C.; Simpson, A.C. Threat intelligence quality dimensions for research and practice. *Digit. Threat. Res. Pract.* **2022**, *3*, 44. <https://doi.org/10.1145/3484202>.
 34. Vielberth, M.; Böhm, F.; Fichtinger, I.; Pernul, G. Security operations center: A systematic study and open challenges. *IEEE Access* **2020**, *8*, 227756–227779. <https://doi.org/10.1109/ACCESS.2020.3045514>.
 35. Alahmadi, B.A.; Axon, L.; Martinovic, I. 99% false positives: A qualitative study of SOC analysts’ perspectives on security alarms. In *31st USENIX Security Symposium (USENIX Security 22)*; USENIX Association: Berkeley, CA, USA, 2022. Available online: <https://www.usenix.org/conference/usenixsecurity22/presentation/alahmadi> (accessed on 27 August 2026).
 36. Iyengar, J.; Thomson, M. QUIC: A UDP-Based Multiplexed and Secure Transport; RFC 9000; Internet Engineering Task Force: Wilmington, DE, USA, 2021. <https://doi.org/10.17487/RFC9000>.
 37. Karagiannis, T.; Molle, M.; Faloutsos, M.; Broido, A. A nonstationary Poisson view of Internet traffic. In *IEEE INFOCOM 2004*; IEEE: New York, NY, USA, 2004; Volume 3, pp. 1558–1569. <https://doi.org/10.1109/INFCOM.2004.1354569>.
 38. Leland, W.E.; Taqqu, M.S.; Willinger, W.; Wilson, D.V. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Trans. Netw.* **1994**, *2*, 1–15. <https://doi.org/10.1109/90.282603>.
 39. Paxson, V.; Floyd, S. Wide area traffic: The failure of Poisson modeling. *IEEE/ACM Trans. Netw.* **1995**, *3*, 226–244. <https://doi.org/10.1109/90.392383>.

40. Clopper, C.J.; Pearson, E.S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **1934**, *26*, 404–413. <https://doi.org/10.1093/biomet/26.4.404>.
41. Rao, J.N.K.; Scott, A.J. A simple method for the analysis of clustered binary data. *Biometrics* **1992**, *48*, 577–585. <https://doi.org/10.2307/2532311>.
42. Shi, W.; Xu, W.; Huang, Y.; Yao, J.; Hu, W.; Wang, D. Flexible-duplex cell-free architecture for secure uplink communications in low-altitude wireless networks. *IEEE J. Sel. Top. Signal Process.* **2026**, *early access*. <https://doi.org/10.1109/JSTSP.2026.3691298>.
43. Shi, W.; Xu, W.; You, X.; Zhao, C.; Wei, K. Intelligent reflection enabling technologies for integrated and green Internet-of-Everything beyond 5G: Communication, sensing, and security. *IEEE Wirel. Commun.* **2023**, *30*, 147–154. <https://doi.org/10.1109/MWC.018.2100717>.
44. Hayes, A.F.; Krippendorff, K. Answering the call for a standard reliability measure for coding data. *Commun. Methods Meas.* **2007**, *1*, 77–89. <https://doi.org/10.1080/19312450709336664>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.