Sarcouncil Journal of Engineering and Computer Sciences

ISSN(Online): 2945-3585

Volume- 05| Issue- 04| 2026

Review Article **Received:** 10-03-2026 | **Accepted:** 05-04-2026 | **Published:** 20-04-2026

Hospital Cyber Attack Forecasting: A Review of Time-Series Methods Used to Predict Security Incidents

Seth Yao Alornyo

School of Electrical Engineering and Computer Science, University of North Dakota, Grand Forks, ND 58202-7165

Abstract: Cyber-attacks on hospitals continue to escalate in sophistication and frequency, threatening patient safety, clinical continuity, and data integrity. This study provides a structured review of time-series forecasting methods for hospital cybersecurity, emphasizing their value in anticipating attacks rather than responding after compromise. It examines statistical, machine learning, deep learning, and hybrid models, assessing their suitability against hospital specific challenges such as sparse and bursty incident data, high non-stationarity, and strict privacy regulations. The review analyzes internal data sources, including incident logs, IDS/IPS alerts, SIEM outputs, and network telemetry, alongside external threat intelligence used as exogenous predictors. Key limitations involve inconsistent labeling, limited dataset size, and evolving attacker behavior. The study concludes that forecasting can significantly enhance preparedness and resilience in healthcare, provided models are healthcare-tailored, context-aware, interpretable, and supported by privacy-preserving data-sharing mechanisms.

Keywords: Cybersecurity, Ransomware, Phishing, Time-series analysis.

INTRODUCTION

Cybersecurity threats against hospitals are rising in frequency and severity, with far-reaching impacts on patient care, operational continuity, and patient privacy. Healthcare institutions have become prime targets for cybercriminals because they hold highly sensitive data and rely on complex, interconnected systems that are often outdated or under-protected (Ahmed *et al*., 2022; Argaw *et al*., 2020). Ransomware attacks alone have surged in recent years and have disrupted clinical operations, forced delays in care delivery, and exposed patient records to extortion (Roumani & Roumani, 2025). These incidents not only incur financial losses but also risk patient safety when critical systems become unavailable. While much work in healthcare cybersecurity has centered on detecting active attacks or preventing breaches, there is a growing recognition that reactive approaches are insufficient. Detection systems identify threats only after they begin, often leaving limited time to respond before damage occurs (Neri *et al*., 2025). Forecasting attacks, on the other hand, aims to anticipate future incidents based on patterns in historical security data. This proactive perspective enables hospitals to allocate defensive resources in advance, adjust staffing in security operations centers, and inform policy and risk mitigation planning with forward-looking insights (Filani *et al*., 2026).

Time-series analysis is a key tool for this predictive approach because most hospital security data are inherently temporal. Incident logs, alerts from intrusion detection systems, and network traffic records are recorded over time, and changes in these data streams can reveal trends and recurring patterns that precede major security incidents. Time-series models can capture these temporal structures, making it possible to identify both long-term trends and short-term anomalies that signal increased risk of attack (Javed *et al*., 2022). Forecasting approaches like ARIMA (Autoregressive Integrated Moving Average), exponential smoothing, and machine learning-based models have already shown promise in capturing the trajectory of cyber threats in broader contexts, including ransomware trends across critical sectors (Roumani & Roumani, 2025). Despite this potential, current surveys in cybersecurity analytics rarely target the specific needs of hospitals. Existing reviews often focus on general intrusion detection or anomaly detection across industries, with limited synthesis of how time-series forecasting methods perform with healthcare data and under healthcare operational constraints (Filani *et al*., 2026; Qureshi & Koo., 2026). Hospitals differ from other sectors because of their unique mix of legacy and modern systems, the criticality of continuous availability, and the direct link between system failures and patient outcomes. This means that forecasting models must not only be accurate but robust to irregular and sparse attack data and capable of producing actionable forecasts for operational decision-making (Khallaf *et al*., 2025).

This review aims to fill that gap. It examines the time-series methods used to forecast cyber attacks

***Corresponding Author:** Seth Yao Alornyo

DOI- https://doi.org/10.5281/zenodo.19661659

in hospital environments, categorizing them by data types, forecasting horizons, and underlying modeling approaches. It also highlights the practical challenges of applying these methods in healthcare settings and identifies areas needing further research. The goal is to provide a comprehensive resource that helps advance proactive cybersecurity planning in hospitals by integrating temporal forecasting into their defensive strategies.

HOSPITAL CYBER SECURITY THREAT LANDSCAPE

Hospitals occupy a unique position in the cyber threat ecosystem because they combine high-value data, mission-critical operations, and complex technological infrastructures. Unlike many other sectors, healthcare organizations cannot tolerate prolonged system outages without risking patient harm (Neri *et al*.,2025). This combination has made hospitals increasingly attractive targets for cybercriminals and state-sponsored actors, while simultaneously limiting the defensive measures that can be deployed without disrupting care delivery (Filani *et al*., 2026).

Common Cyberattacks Targeting Hospitals

Ransomware represents the most disruptive and widely reported cyber threat to hospitals in recent years. These attacks encrypt clinical and administrative systems, rendering electronic health records, imaging platforms, and scheduling systems inaccessible (Argaw *et al*., 2020). Studies have shown that ransomware incidents in healthcare lead to emergency department diversions, delayed procedures, and increased patient risk, particularly when attacks coincide with peak service demand (Bonsu & Opoku., 2025; Brantly, 2025). Unlike other sectors where downtime primarily results in financial loss, ransomware in hospitals directly affects patient outcomes, making it a public safety issue rather than a purely technical one.

Phishing and credential theft continue to serve as primary entry points for attackers. Healthcare staff frequently interact with emails and digital systems under time pressure, which increases susceptibility to social engineering attacks. Once credentials are compromised, attackers can move laterally across hospital networks, escalate privileges, and deploy more advanced attacks such as ransomware or data exfiltration. These attacks exploit human factors rather than technical vulnerabilities, making them difficult to eliminate entirely through traditional security controls (Cabello *et al*., 2020). Distributed denial-of-service attacks have also been observed in healthcare environments, particularly against hospital websites, telemedicine platforms, and patient portals. While DDoS (Distributed Denial of Service) attacks do not typically compromise data integrity, they can disrupt access to critical online services and delay communication between patients and providers. Research indicates that even short service interruptions can have cascading effects in hospitals where systems are tightly coupled and interdependent (Coventry & Branley, 2018). Insider threats pose a distinct challenge in healthcare because of the large number of users who legitimately require access to sensitive systems and data. Insider incidents may be malicious, such as intentional data theft, or unintentional, such as accidental disclosure or unsafe handling of credentials. Studies show that insider threats in healthcare are often linked to workflow pressures, insufficient training, and overly broad access permissions, all of which increase the attack surface without clear malicious intent (Jalali *et al*., 2019).

The increasing deployment of connected medical devices has introduced new attack vectors through medical device exploitation. Devices such as infusion pumps, patient monitors, and imaging systems often run on legacy software and are difficult to patch without interrupting clinical use. Vulnerabilities in these devices can allow attackers to gain persistent access to hospital networks or interfere with device functionality, raising concerns about both cybersecurity and physical patient safety (Coventry & Branley, 2018).

Unique Characteristics of Hospital Security Data

Hospital security data differs significantly from that of other sectors due to its scale, diversity, and operational constraints. Hospitals generate extremely high volumes of security alerts from firewalls, intrusion detection systems, endpoint protection tools, and network monitoring platforms (Cabello *et al*., 2020). However, a large proportion of these alerts are low severity or false positives, making it difficult for security teams to identify meaningful patterns without advanced analytical methods (Argaw *et al*., 2020).

The heterogeneity of hospital systems further complicates security data analysis. Hospital environments integrate traditional information technology systems, operational technology such as building management and power systems, and Internet of Medical Things devices embedded in

Publisher: SARC Publisher

clinical workflows (Bonsu & Opoku., 2025). These systems produce data in different formats, at different temporal resolutions, and with varying levels of reliability. Time-series modeling in this context must account for irregular sampling, missing data, and abrupt changes caused by system upgrades or clinical emergencies.

Strict uptime and safety requirements also shape the nature of hospital security data. Many systems cannot be taken offline for patching or reconfiguration without disrupting care, leading to prolonged exposure to known vulnerabilities (Brantly, 2025). As a result, attack patterns in hospitals often exhibit persistence and recurrence rather than isolated events, creating temporal dependencies that are well suited to time-series analysis but poorly handled by static security assessments (Jalali *et al*., 2019).

Regulatory pressure further influences how hospital security data can be collected, stored, and analyzed. Regulations such as HIPAA (Health Insurance Portability and Accountability Act) and GDPR (General Data Protection Regulation) impose strict controls on access to patient data and audit logs, limiting data sharing and the creation of large public datasets. While these regulations are essential for protecting patient privacy, they also constrain cybersecurity research and model validation, contributing to fragmented and institution-specific forecasting approaches (Cabello *et al*., 2020).

Why Forecasting Is Critical in Healthcare

Forecasting cyber attacks is particularly critical in healthcare because of the sector's limited tolerance for disruption and its reliance on proactive planning. Reactive security measures, which respond only after an attack has begun, often fail to prevent operational shutdowns in hospitals. Forecasting enables healthcare organizations to anticipate periods of elevated risk and implement targeted defenses before systems are compromised (Neri *et al*., 2025).

From a patient safety perspective, forecasting supports continuity of care by reducing the likelihood of sudden system outages. Predictive insights can inform decisions such as postponing non-urgent system changes during high-risk periods or increasing monitoring of critical devices when attack likelihood is elevated (Bonsu & Opoku., 2025). Research has shown that hospitals with more proactive cybersecurity strategies experience fewer severe disruptions during large-scale cyber incidents (Filani *et al*., 2026).

Forecasting also plays a key role in resource planning and cyber resilience. Security operations centers in hospitals often operate with limited staff and budgets. Time-series forecasts of incident volume can guide staffing decisions, incident response preparedness, and investment in defensive technologies. Over time, these predictive capabilities contribute to institutional resilience by enabling hospitals to adapt to evolving threat landscapes rather than reacting to each incident in isolation.

FUNDAMENTALS OF TIME-SERIES ANALYSIS FOR CYBER ATTACK FORECASTING

Time-series analysis provides the theoretical foundation for forecasting cyber attacks because security events in hospital environments are recorded sequentially over time. Logs from intrusion detection systems, firewall alerts, authentication records, and incident reports all reflect evolving system states and attacker behavior (Bonsu & Opoku., 2025). Unlike static analysis methods that examine isolated observations, time-series approaches explicitly model temporal dependence, making them well suited for capturing trends, recurring patterns, and changes in attack dynamics. This section outlines the core time-series concepts and properties relevant to hospital cyber security and clarifies the distinction between forecasting and anomaly detection, which is central to the scope of this review.

Time-Series Concepts

A time series consists of observations collected at successive time intervals, where the ordering of data points carries meaningful information. In hospital cyber security, time series may represent counts of security incidents per hour, volume of suspicious network traffic per day, or frequency of ransomware alerts per week. These series can be either univariate or multivariate. A univariate time series models a single variable, such as the daily number of detected phishing attempts, and is often used when data availability is limited or interpretability is prioritized. Multivariate time series incorporate multiple related variables, such as combining incident counts with network traffic metrics or vulnerability disclosures, allowing models to capture interactions between different aspects of the security environment (Coventry & Branley).

Publisher: SARC Publisher

Hospital security data may also be discrete or continuous in nature. Discrete time series are common in cyber security and include event counts or binary indicators of attack occurrence within fixed time windows. Continuous time series arise when measuring variables such as bandwidth usage or latency, which can fluctuate continuously over time. The choice between discrete and continuous representations affects model selection and aggregation strategies, particularly in hospitals where raw data are often irregular and must be summarized into consistent intervals for analysis (Ahmed *et al*., 2022).

Forecasting horizons further distinguish time-series applications. Short-term forecasting focuses on predicting events in the near future, such as estimating the likelihood of attack spikes in the next few hours or days. These forecasts are valuable for operational decision-making, including staffing security operations centers or increasing monitoring during anticipated high-risk periods. Long-term forecasting aims to capture broader trends over weeks, months, or years, supporting strategic planning, budg*et al*location, and policy development. In healthcare, both horizons are important, but they pose different challenges due to the variability and rarity of severe cyber incidents (Roumani & Roumani, 2025).

Core Time-Series Properties in Security Data

Security-related time series in hospitals exhibit several characteristic properties that influence forecasting performance. One fundamental property is trend, which reflects long-term increases or decreases in attack activity. Studies have documented upward trends in ransomware and phishing incidents in healthcare, driven by digital transformation and increasing attacker sophistication (Neri *et al*., 2025; Argaw *et al*., 2020). Accurately modeling trends is essential for understanding whether observed changes represent temporary fluctuations or sustained shifts in the threat landscape.

Seasonality is another prominent feature, where attack frequency follows regular cycles. In hospital environments, seasonal patterns may align with weekdays versus weekends, staffing schedules, or known attacker behaviors that exploit reduced monitoring during off-peak hours. For example, some studies have observed higher attack activity during weekends or holidays when response capacity is lower, creating predictable temporal structures that can be exploited by forecasting models (Jain *et al*., 2024).

Burstiness and sparsity are particularly pronounced in hospital cyber security data. Many attack types occur infrequently but in intense bursts, such as coordinated ransomware campaigns that target multiple hospitals within a short time frame. Between these bursts, incident counts may remain near zero for extended periods. This irregularity challenges traditional time-series models that assume relatively stable variance and continuous activity, and it increases the risk of false confidence in forecasts if rare but high-impact events are not properly accounted for (Roumani & Roumani, 2025).

Non-stationarity further complicates modeling efforts. A time series is non-stationary when its statistical properties, such as mean and variance, change over time. In healthcare cyber security, non-stationarity arises from evolving attacker tactics, system upgrades, policy changes, and external events such as public vulnerability disclosures. These changes can render historical patterns less informative for future prediction, requiring models that adapt to concept drift and structural breaks rather than assuming stable behavior (Ahmed *et al*., 2022).

Forecasting Versus Anomaly Detection

Although forecasting and anomaly detection are often discussed together in cyber security research, they serve distinct purposes and rely on different assumptions. Anomaly detection focuses on identifying deviations from expected or normal behavior, flagging unusual events that may indicate an ongoing or previously unknown attack (Nnadi & Opoku., 2025). These methods are typically reactive, as they operate on current or past data to detect abnormalities once they occur. In hospital environments, anomaly detection is widely used for real-time monitoring but can generate large numbers of false positives due to the complexity and variability of clinical systems (Argaw *et al*., 2020).

Forecasting, in contrast, aims to predict future attack occurrences or activity levels based on historical temporal patterns. Rather than identifying unexpected deviations, forecasting estimates what is likely to happen next, providing a probabilistic view of future risk. This forward-looking perspective is particularly valuable in healthcare, where proactive preparation can prevent operational shutdowns and reduce patient

Publisher: SARC Publisher

harm. Forecasting supports decisions such as when to reinforce defenses, schedule maintenance, or increase staffing before an attack materializes (Roumani & Roumani, 2025).

This paper focuses on forecasting rather than anomaly detection because the healthcare sector requires anticipatory strategies that extend beyond real-time alerting. While anomaly detection remains essential for identifying active threats, forecasting complements these systems by enabling hospitals to move from reactive defense to proactive resilience. By synthesizing time-series forecasting methods within the hospital cyber security context, this review emphasizes approaches that support planning, risk mitigation, and long-term cyber resilience rather than solely immediate threat detection.

DATA SOURCES FOR HOSPITAL CYBER ATTACK FORECASTING

Effective forecasting of cyber attacks in hospital environments depends fundamentally on the quality, diversity, and temporal resolution of available data. Unlike traditional enterprise settings, hospitals generate security-relevant data from clinical, administrative, and operational systems that operate continuously and under strict safety constraints. Time-series forecasting models rely on these data streams to learn temporal patterns associated with attack emergence, escalation, and recurrence. This section reviews the primary internal and external data sources used in hospital cyber attack forecasting and discusses preprocessing and feature engineering strategies necessary to make these data suitable for time-series analysis.

Internal Hospital Data Sources

Internal hospital data form the core input for most cyber attack forecasting models because they directly reflect the operational security state of healthcare systems. Security incident logs are among the most commonly used data sources and typically record confirmed or suspected security events, including malware detections, unauthorized access attempts, and policy violations (Nnadi & Opoku., 2025). These logs are often curated by security teams and provide labeled historical records that are valuable for supervised forecasting tasks, although they may underrepresent low-level or undetected activity (Argaw *et al*., 2020).

Intrusion detection and prevention system alerts provide higher-frequency data streams that capture suspicious behavior at the network or host level. IDS (Intrusion Detection System) and IPS (Intrusion Prevention System) alerts are particularly useful for short-term forecasting because they offer granular temporal resolution, often recorded in seconds or minutes (Cabello *et al*., 2020). However, these alerts are known to be noisy and prone to false positives, which can obscure true attack patterns unless properly aggregated and filtered (Landauer, 2025). In hospital environments, where clinical systems generate atypical traffic patterns, distinguishing benign anomalies from malicious behavior is especially challenging.

Firewall and security information and event management logs represent another critical internal data source. Firewalls record connection attempts, blocked traffic, and policy violations, while SIEM (Security Information and Event Management) platforms aggregate logs across multiple systems, providing a centralized temporal view of security activity. These data sources are well suited for multivariate time-series modeling because they capture correlated activity across systems and departments, which is common during coordinated attacks such as ransomware campaigns (Ahmed *et al*., 2022).

Network traffic metrics, including packet counts, flow volumes, and bandwidth utilization, provide continuous time-series data that can reveal early indicators of attack preparation or lateral movement (Nnadi & Opoku., 2025). In hospitals, network traffic data often reflect interactions between IT systems, operational technology, and Internet of Medical Things devices, creating complex temporal patterns that differ from those observed in other sectors. These metrics are particularly useful for forecasting distributed denial-of-service activity and large-scale scanning behavior.

Table 1. Internal hospital data sources for cyberattack forecasting

Data source	Temporal resolution	Typical use in forecasting	Key limitations
Security incident logs	Daily to weekly	Long-term trend forecasting	Underreporting, delayed labeling
IDS/IPS alerts	Seconds to minutes	Short-term spike	High false-positive rates

Publisher: SARC Publisher

		prediction	
Firewall and SIEM logs	Minutes to hours	Multivariate forecasting	Data heterogeneity
Network traffic metrics	Continuous	Early warning indicators	High dimensionality

Table 2. External and contextual data sources used in hospital cyber attack forecasting

Data source	Forecasting role	Integration approach
Threat intelligence feeds	Anticipate coordinated attacks	Exogenous variables
CVE timelines	Predict post-disclosure exploitation	Event-based features
Public breach reports	Validate long-term trends	Aggregated benchmarks
News and OSINT	Contextual risk signals	Temporal alignment

Raw security data from hospital environments are rarely suitable for direct input into time-series forecasting models. Preprocessing is required to transform heterogeneous, irregular data into structured temporal representations. Event aggregation is a foundational step in which raw events are summarized into fixed time windows, such as hourly or daily counts. The choice of aggregation window directly affects model sensitivity; shorter windows capture rapid attack dynamics, while longer windows smooth noise and highlight trends (Roumani & Roumani, 2025).

Lag features are widely used to encode temporal dependence by including past values of a variable as predictors of future behavior. In hospital cyber security, lagged incident counts or alert volumes can capture attacker persistence and recurring attack cycles. Rolling statistics, such as moving averages and rolling variances, further summarize recent behavior and help models distinguish between baseline activity and emerging threats, particularly in bursty data environments (Jain *et al*., 2024). Cyclical time features encode periodic patterns inherent in hospital operations and attacker behavior. Features representing hour of day, day of week, or holiday periods allow models to learn systematic variations, such as increased attack activity during weekends or night shifts when monitoring may be reduced. These features are especially important in healthcare settings, where staffing and system usage follow predictable cycles that interact with security risk (Argaw *et al*., 2020).

TIME-SERIES FORECASTING METHODS USED IN HOSPITAL CYBER SECURITY

Time-series forecasting methods used in hospital cyber security span classical statistical approaches, machine learning models, deep learning architectures, and hybrid systems. Each category differs in terms of assumptions about temporal structure, data requirements, interpretability, and computational cost. In hospital environments, model choice is influenced not only by predictive accuracy but also by data sparsity, non-stationarity, regulatory constraints, and the need for operational transparency. This section synthesizes the principal forecasting approaches applied in cybersecurity research and evaluates their suitability for hospital settings.

Statistical Forecasting Models

Statistical time-series models remain widely used in cyber attack forecasting due to their interpretability and modest data requirements. Autoregressive Integrated Moving Average models, including seasonal extensions such as SARIMA (Seasonal Autoregressive Integrated Moving Average), are among the most established approaches. ARIMA models capture temporal dependence by expressing current values as a function of past observations and past forecast errors, while differencing operations address non-stationarity. SARIMA extends this framework to incorporate seasonal components, which is particularly relevant in cybersecurity where weekly or monthly attack cycles are common (Landauer, 2025).

In hospital cyber security contexts, ARIMA-type models are often applied to incident count data, ransomware trends, or alert volumes aggregated over fixed intervals. Their strength lies in modeling structured patterns such as gradual upward trends in phishing attempts or recurring spikes during weekends. Studies have shown that ARIMA-based models can provide competitive short-term forecasts when attack behavior follows relatively stable temporal patterns (Roumani & Roumani, 2025). Moreover, the transparency of model parameters supports interpretability, which is important in regulated healthcare environments.

Exponential smoothing methods, including Error-Trend-Seasonal models and Exponentially

Publisher: SARC Publisher

Weighted Moving Averages, offer an alternative statistical approach. These methods assign greater weight to recent observations, allowing forecasts to adapt more quickly to changes. In hospital security monitoring, EWMA (Exponentially Weighted Moving Average) is often used to track deviations in alert volumes or network traffic baselines. While these models are computationally efficient and easy to implement, they may struggle with abrupt structural breaks, such as sudden ransomware outbreaks that do not follow prior trends.

Despite their advantages, statistical models face limitations in hospital data environments characterized by burstiness, sparsity, and high-dimensional multivariate inputs. They typically assume linear relationships and may underperform when attack patterns are influenced by complex interactions among multiple systems or external threat signals (Ahmed *et al*., 2022). As hospital infrastructures grow more interconnected, purely linear models may be insufficient to capture evolving attacker strategies.

Machine Learning-Based Models

Machine learning approaches relax many assumptions imposed by statistical models and can capture nonlinear relationships between variables. Support Vector Regression has been used in cyber forecasting tasks to model complex temporal dependencies while maintaining robustness to noise. SVR (Support Vector Regression) performs well when the dataset is moderate in size and when feature engineering captures relevant lagged and contextual variables. In hospital settings, SVR can integrate incident counts with vulnerability disclosures or threat intelligence signals to improve predictive accuracy (Landauer, 2025). Random Forests represent another widely applied method. By constructing ensembles of decision trees, Random Forests can model nonlinear interactions and handle heterogeneous input features. In cyber security forecasting, they have been used to predict incident spikes and ransomware activity based on historical trends and contextual indicators (Roumani & Roumani, 2025). Their ability to estimate feature importance provides partial interpretability, which is beneficial in healthcare decision-making contexts.

k-Nearest Neighbors (kNN) forecasting methods rely on identifying historical periods with similar patterns and projecting future behavior based on those analogues. While simple to implement, kNN approaches can struggle with high-dimensional data and may be sensitive to noise. In hospital cyber environments where data may be sparse or irregular, kNN performance depends heavily on effective preprocessing and similarity metrics. Machine learning models generally outperform traditional statistical models when attack dynamics are nonlinear and influenced by multiple interacting variables. However, they require larger datasets and careful hyperparameter tuning. In hospitals, data privacy constraints and limited access to large, labeled datasets may restrict the effectiveness of purely data-driven approaches (Argaw *et al*., 2020).

Table 3. Machine learning forecasting models in hospital cybersecurity

Model	Key strength	Key limitation	Data requirement
SVR	Robust to noise, nonlinear modeling	Sensitive to parameter choice	Moderate
Random Forest	Handles heterogeneous features	Less effective for long-term memory	Moderate to high
kNN	Conceptually simple	Sensitive to noise and dimensionality	Moderate

Deep Learning Approaches

Deep learning architectures have gained prominence in cyber attack forecasting due to their ability to model long-range dependencies and complex temporal dynamics. Long Short-Term Memory networks are particularly suited for sequential data because they maintain memory cells that preserve information across long time horizons. In cybersecurity research, LSTM (Long Short-Term Memory Network) models have demonstrated improved performance over ARIMA and classical machine learning in predicting attack frequencies and intrusion patterns (Landauer, 2025). Gated Recurrent Units offer a simplified alternative to LSTM with fewer parameters, reducing computational overhead while retaining the ability to capture temporal dependencies. In hospital settings where computational resources may be limited and real-time forecasting is required, GRU (Gated Recurrent Unit) models

Publisher: SARC Publisher

may provide a balance between complexity and performance.

Temporal Convolutional Networks (TCN) apply convolutional filters across time steps, enabling parallel computation and efficient training. These models can capture both short- and long-range dependencies without the vanishing gradient issues associated with recurrent networks. Transformer-based architectures further extend this capability through attention mechanisms that dynamically weight the importance of different time steps. Although originally developed for language modeling, transformers have been increasingly applied to time-series forecasting tasks due to their scalability and flexibility.

Deep learning models are particularly advantageous when large volumes of high-resolution data are available, such as continuous network traffic streams. However, their performance depends on substantial training data and computational resources. In hospital contexts, where incident data may be sparse and privacy restrictions limit data sharing, overfitting and generalization remain concerns (Ahmed *et al*., 2022).

Table 4. Deep learning models for hospital cyber attack forecasting

Model	Advantage	Limitation	Suitability
LSTM	Captures long-term dependencies	High computational cost	Complex temporal patterns
GRU	Efficient recurrent modeling	Slightly reduced representational power	Resource-constrained settings
Temporal CNN	Parallel processing, stable gradients	Requires careful architectural design	High-frequency data
Transformer	Attention-based long-range modeling	Data-hungry and complex	Large-scale multivariate data

Hybrid and Ensemble Models

Hybrid and ensemble models combine complementary strengths of statistical and machine learning approaches. One common strategy integrates ARIMA models for linear trend components with neural networks to capture nonlinear residual structures. Such hybrid systems have been shown to outperform standalone models in ransomware trend forecasting and other cyber security applications (Roumani & Roumani, 2025).

Ensemble approaches aggregate predictions from multiple models to improve robustness and reduce variance. In hospital cyber security, ensembles can combine short-term and long-term models or integrate internal incident data with external threat intelligence feeds (Johnson., 2022). Multi-source forecasting systems that incorporate vulnerability disclosures, news signals, and internal alert data represent a promising direction for improving resilience against coordinated attacks (Khallaf *et al*., 2025).

Hybrid approaches are particularly attractive in healthcare because they allow interpretable statistical components to coexist with high-capacity neural models. This balance can support regulatory transparency while maintaining predictive performance. However, increased complexity may hinder deployment in resource-constrained hospital IT departments.

COMPARATIVE ANALYSIS OF FORECASTING APPROACHES

The forecasting methods discussed in the previous section differ substantially in their assumptions, performance characteristics, and operational requirements. In hospital environments, model selection cannot rely solely on predictive accuracy. Instead, it must consider interpretability, data availability, computational constraints, and deployment feasibility within regulated and safety-critical systems. This section synthesizes the reviewed approaches using consistent evaluation criteria and assesses their suitability for hospital cyber security forecasting.

Comparison Criteria

Prediction accuracy is the most commonly reported performance metric in forecasting studies. Error-based measures such as mean absolute error, root mean square error, and mean absolute percentage error are frequently used when forecasting incident counts or attack volumes (Javed *et al*., 2022). In cybersecurity contexts where forecasting may be framed as predicting whether an attack spike will occur, classification-oriented metrics such as precision, recall, and F1-score are also used (Roumani & Roumani, 2025). Deep learning models, particularly LSTM and transformer-based architectures, often demonstrate superior predictive accuracy when sufficient

Publisher: SARC Publisher

historical data are available. However, performance gains may diminish in hospital environments characterized by sparse or irregular incident data.

Interpretability is particularly important in healthcare settings, where decisions based on model outputs may influence staffing, resource allocation, and operational planning. Statistical models such as ARIMA and exponential smoothing provide clear parameter interpretations, allowing analysts to understand how trends and seasonality influence forecasts. Machine learning models like Random Forests offer partial interpretability through feature importance measures. In contrast, deep learning architectures are often described as black-box models, making it more difficult to justify decisions in environments subject to regulatory scrutiny and accountability requirements (Ahmed *et al*., 2022). Hospitals may therefore favor models that balance performance with transparency.

Data requirements vary significantly across forecasting approaches. Statistical models can operate effectively with moderate historical data and are relatively robust when datasets are small. Machine learning and deep learning methods typically require larger volumes of labeled data to avoid overfitting and to generalize effectively (Johnson., 2022). In hospital cyber security, access to large, high-quality datasets is often constrained by privacy regulations and limited incident reporting. These constraints may reduce the practical feasibility of highly complex models despite their theoretical advantages (Argaw *et al*., 2020).

Computational complexity also influences model selection. Statistical models are computationally efficient and suitable for real-time implementation on standard hospital IT infrastructure. Machine learning models require more processing power but remain manageable in most operational contexts. Deep learning architectures, particularly transformers, demand significant computational resources and specialized hardware for training. While inference can be optimized, initial deployment and maintenance costs may pose challenges for smaller hospitals with limited cybersecurity budgets.

Short-Term Versus Long-Term Forecasting

Forecasting objectives in hospital cyber security vary depending on the time horizon. Short-term forecasting focuses on anticipating immediate spikes in attack activity, such as sudden increases in phishing attempts or coordinated ransomware campaigns. These predictions support operational responses, including temporary reinforcement of monitoring, patch management prioritization, or incident response readiness (Johnson., 2022). Models that emphasize recent trends and short memory, such as exponential smoothing and certain machine learning approaches, often perform well in short-term contexts because they adapt quickly to new patterns (Qureshi & Koo., 2026).

Long-term forecasting addresses broader trends in cyber threat evolution. Hospitals may use long-range forecasts to inform strategic planning, budg*et al*location, and infrastructure investments. ARIMA and SARIMA models are frequently applied to capture persistent upward or downward trends and seasonal cycles over extended periods. Deep learning models can also support long-term forecasting when trained on sufficiently large datasets, but their advantage depends on stable underlying patterns and consistent data collection.

The distinction between short-term spike prediction and long-term trend forecasting is critical in healthcare. Short-term forecasts directly influence patient safety and operational continuity, while long-term forecasts shape institutional resilience strategies. Selecting the appropriate modeling approach requires alignment between forecasting horizon and organizational objectives (Bajrić., 2020).

Suitability for Hospital Environments

Forecasting models must be evaluated not only on analytical performance but also on practical suitability within hospital environments. Real-time applicability is essential when forecasts are used to support active monitoring and decision-making. Statistical and lightweight machine learning models are often easier to integrate into existing security information and event management systems due to their computational efficiency and simpler deployment requirements (Argaw *et al*., 2020). Deep learning systems may require dedicated infrastructure and maintenance expertise that not all hospitals possess.

Robustness to missing or incomplete data is another critical factor. Hospital security logs may contain gaps due to system upgrades, device failures, or data retention policies. Models that assume continuous and complete time series may perform poorly when confronted with such

Publisher: SARC Publisher

irregularities. Techniques that incorporate rolling statistics or adaptive weighting can mitigate some of these effects, but complex neural models may still be sensitive to inconsistent input streams (Jain *et al*., 2024).

Deployment constraints further shape model suitability. Hospitals operate under strict regulatory frameworks and must ensure that predictive systems do not compromise patient privacy or system availability. Models that require centralized aggregation of sensitive logs may encounter compliance challenges, particularly under data protection regulations. In addition, cybersecurity teams in hospitals are often smaller than those in large enterprises, limiting capacity for maintaining complex forecasting pipelines (Johnson., 2022).

Overall, while deep learning and hybrid models may offer higher predictive accuracy in ideal conditions, statistical and carefully tuned machine learning approaches often provide a more balanced solution for hospital environments. The optimal approach depends on institutional size, data availability, regulatory context, and operational priorities. A key insight emerging from this comparison is that forecasting in hospitals requires not only methodological sophistication but also practical alignment with healthcare system constraints.

CHALLENGES AND LIMITATIONS

Despite advances in time-series forecasting methods, applying these techniques to hospital cyber security remains constrained by structural, methodological, and evaluative challenges. Many of these limitations stem from the unique operational and regulatory environment of healthcare (Salama & Al-Turjman., 2024). While forecasting models show promise in controlled research settings, their real-world deployment in hospitals is complicated by data scarcity, evolving attack behavior, and the absence of standardized evaluation frameworks. This section critically examines these challenges, drawing on recent cybersecurity and healthcare informatics research (Qureshi & Koo., 2026).

Data Challenges

One of the most significant barriers to effective forecasting in hospital cyber security is data sparsity. Although hospitals generate large volumes of raw logs, confirmed cyber attack incidents are relatively rare events when aggregated at daily or weekly resolution. Severe ransomware outbreaks or coordinated attacks occur infrequently but have substantial impact (Salama & Al-Turjman., 2024). This leads to time series characterized by long periods of low activity punctuated by sudden spikes. Sparse and bursty data reduce the statistical power of many forecasting models and increase the risk of overfitting, particularly for complex machine learning and deep learning approaches (Bajrić., 2020).

Label scarcity further complicates predictive modeling. In many hospital environments, security incidents are not consistently labeled with high confidence due to resource constraints and the difficulty of forensic investigation. Some malicious activities may remain undetected, while others may be misclassified. Supervised forecasting models depend on accurate historical labels to learn meaningful patterns. When labels are incomplete or inconsistent, predictive accuracy deteriorates, and model validation becomes unreliable (Argaw *et al*., 2020). This problem is particularly acute in healthcare institutions where cybersecurity teams are often small relative to the scale of digital infrastructure.

Privacy and access restrictions represent another structural limitation. Regulations such as HIPAA and GDPR impose strict controls on how patient-related data and system logs can be accessed, shared, and analyzed. While these regulations are essential for protecting patient confidentiality, they limit the availability of large, centralized datasets that could support robust forecasting research (Pavlova & Albert., 2025). Many hospitals are reluctant to share detailed incident logs, even in anonymized form, due to reputational and legal concerns. As a result, most forecasting studies rely on institution-specific datasets, reducing generalizability across healthcare systems (Ahmed *et al*., 2022).

Modeling Challenges

Hospital cyber attack patterns are inherently non-stationary. Non-stationarity arises when statistical properties such as mean, variance, or autocorrelation change over time. In healthcare cyber security, these changes may result from new attack techniques, software upgrades, infrastructure modernization, or external geopolitical events that alter threat intensity (Salama & Al-Turjman., 2024). Traditional statistical models often assume stable temporal behavior, and their performance degrades when these assumptions are violated (Bajrić., 2020).

Publisher: SARC Publisher

While adaptive or rolling-window techniques can partially address non-stationarity, abrupt structural shifts remain difficult to model.

Closely related to non-stationarity is the phenomenon of concept drift. Concept drift occurs when the underlying relationship between input variables and attack outcomes evolves over time. For example, vulnerability disclosures that previously had little predictive value may become highly relevant if attackers begin systematically exploiting them (Salama & Al-Turjman., 2024). In hospital environments undergoing rapid digital transformation, such shifts can occur frequently. Machine learning models trained on historical data may fail to generalize to new threat conditions unless retrained or adapted dynamically (Ahmed *et al*., 2022).

Rare but high-impact events pose an additional modeling challenge. Ransomware attacks targeting hospitals may be infrequent but catastrophic. Forecasting models trained primarily on normal or low-level activity may struggle to anticipate these extreme events because they lack sufficient historical examples (Pavlova & Albert., 2025). Statistical models may underestimate the probability of extreme spikes, while deep learning models may overfit limited instances. This imbalance between frequency and impact creates tension between optimizing average predictive accuracy and ensuring preparedness for worst-case scenarios (Argaw *et al*., 2020).

Evaluation Challenges

Beyond data and modeling constraints, evaluation practices in hospital cyber forecasting remain inconsistent. A major limitation is the lack of benchmark healthcare-specific datasets. In other domains such as image recognition or natural language processing, standardized datasets enable direct comparison across methods. In hospital cyber security, no widely accepted public dataset exists that captures longitudinal attack data under realistic operational conditions. Consequently, studies often rely on proprietary or simulated datasets, limiting reproducibility and cross-study comparability (Javed *et al*., 2022).

Inconsistent evaluation metrics further hinder meaningful comparison. Some studies report error-based measures such as mean absolute error or root mean square error, while others use classification-oriented metrics such as precision and recall when forecasting is framed as a binary spike prediction task (Bajrić., 2020). Differences in aggregation windows, forecasting horizons, and validation strategies make it difficult to determine whether performance differences reflect genuine methodological superiority or merely variations in experimental design. Without standardized reporting practices, drawing firm conclusions about optimal forecasting approaches for hospital environments remains challenging (Johnson., 2022).

Taken together, these challenges underscore that forecasting cyber attacks in hospitals is not simply a technical modeling problem. It is constrained by data governance, institutional capacity, regulatory requirements, and the evolving nature of cyber threats (Jain *et al*., 2024). Addressing these limitations will require not only methodological innovation but also collaborative data-sharing frameworks and standardized evaluation protocols tailored to healthcare contexts (Salama & Al-Turjman, 2024).

FUTURE RESEARCH DIRECTIONS

The distinctive operational, regulatory, and safety-critical nature of healthcare systems requires models that move beyond generic cyber forecasting frameworks. One important direction is the development of domain-specific forecasting models designed explicitly for healthcare infrastructure. Most existing time-series forecasting approaches are developed using datasets from enterprise IT networks, financial systems, or general internet traffic (Bajrić., 2020). Hospitals, however, operate hybrid environments that integrate traditional IT systems with operational technology and Internet of Medical Things devices. These systems exhibit different temporal patterns, maintenance cycles, and vulnerability exposure profiles. Domain-specific models could incorporate healthcare-driven temporal signals such as scheduled maintenance windows, device calibration cycles, electronic health record update schedules, and known clinical workflow rhythms (Khallaf *et al*., 2025). Research in healthcare cybersecurity consistently highlights the need for tailored security frameworks rather than direct transplantation of enterprise solutions (Argaw *et al*., 2020). Extending this principle to forecasting models would likely improve predictive reliability and contextual relevance.

Lastly, adding hospital operational indicators as external variables in forecasting models can improve early warning systems (Pavlova & Albert, 2025). Research also shows that cybersecurity in healthcare should be examined together with

Publisher: SARC Publisher

clinical workflows and organizational processes (Jain *et al*., 2024). Including this context may reduce false positives and better predict high-risk periods. Explainable artificial intelligence is a key research priority. Models such as LSTM, GRU, and Transformer perform well in sequence prediction, but their decision processes are often unclear. This limits their use in safety-critical areas like healthcare. Hospital leaders and security teams need clear reasons behind predictions, especially when allocating resources (Pavlova & Albert, 2025). If a model predicts high ransomware risk, it must know which factors drove that result. Methods such as attention visualization and feature attribution can improve transparency in time-series models (Javed *et al*., 2022). Testing these approaches in hospital cyber forecasting can strengthen trust, support compliance, and improve real-world use.

CONCLUSION

This research demonstrates that time-series forecasting offers a promising approach to strengthening proactive cybersecurity in hospitals, where traditional reactive defenses are insufficient to prevent disruptions that compromise patient care. Statistical approaches like ARIMA and exponential smoothing remain valuable because they are interpretable and effective with small or constrained datasets. Meanwhile, advanced machine learning and deep learning models such as SVR, Random Forests, LSTM, GRU, and Transformers achieve superior modeling of complex temporal behaviors but depend on extensive, high-quality data. Hybrid and ensemble approaches present a balanced solution by combining interpretability with predictive power.

However, hospital environments pose persistent challenges, including sparse and bursty attack patterns, inconsistent labeling, strict data-privacy regulations, and rapidly evolving threat landscapes. These constraints limit model performance and hinder the creation of standardized evaluation frameworks. To improve forecasting reliability, future work should focus on developing healthcare-specific models, integrating operational and clinical context, enabling explainable AI for transparency, and adopting collaborative, privacy-preserving data-sharing mechanisms. By addressing these needs, forecasting can evolve into a practical, trustworthy tool that enhances preparedness, guides resource allocation, and strengthens the overall cyber resilience of healthcare systems.

REFERENCES

1. Ahmed, M. A., Sindi, H. F., & Nour, M. "Cybersecurity in hospitals: an evaluation model." *Journal of Cybersecurity and Privacy* 2.4 (2022): 853-861.
2. Argaw, S. T., Troncoso-Pastoriza, J. R., Lacey, D., Florin, M. V., Calcavecchia, F., Anderson, D., & Flahault, A. "Cybersecurity of Hospitals: discussing the challenges and working towards mitigating the risks." *BMC medical informatics and decision making* 20.1 (2020): 146.
3. Argaw, S. T., Troncoso-Pastoriza, J. R., Lacey, D., Florin, M. V., Calcavecchia, F., Anderson, D., & Flahault, A. "Cybersecurity of Hospitals: discussing the challenges and working towards mitigating the risks." *BMC medical informatics and decision making* 20.1 (2020): 146.
4. Bonsu, M. A., & Opoku, J. A. "Privacy Challenges in IoT: Assessing Data Protection Risks and Strategies for Secure User Adoption." (2025).
5. Brantly, N. D. "The US health system vulnerabilities." *BMC Health Services Research* (2025).
6. Cabello, J. C., Karimipour, H., Jahromi, A. N., Dehghantanha, A., & Parizi, R. M. "Big-data and cyber-physical systems in healthcare: Challenges and opportunities." *Handbook of Big Data Privacy* (2020): 255-283.
7. Coventry, L., & Branley, D. "Cybersecurity in healthcare: A narrative review of trends, threats and ways forward." *Maturitas* 113 (2018): 48-52.
8. Filani, A., Addotey, N., & Opoku, J. A. "Cybersecurity Threat Landscape in the US Healthcare Sector: Trends, Risks, and National Implications."
9. Jain, S., Ashok, P., & Prabhu, S. "Emerging technologies for cybersecurity in healthcare: Evaluating risks and implementing standards." *2024 International Conference on Cybernation and Computation (CYBERCOM)*. IEEE, (2024).
10. Javed, A., Lakoju, M., Burnap, P., & Rana, O. "Security analytics for real-time forecasting of cyberattacks." *Software: Practice and Experience* 52.3 (2022): 788-804.
11. Johnson, C. K. "Decision-making biases in cybersecurity: measuring the impact of the sunk cost fallacy to delay attacker behavior." Arizona State University, (2022).
12. Khallaf, F., El-Shafai, W., El-Rabaie, E. S. M., & Abd El-Samie, F. E. "A Systematic Review

Publisher: SARC Publisher

of New Technologies for Cybersecurity Healthcare Applications: A Systematic and Comprehensive Study." *Transactions on Emerging Telecommunications Technologies* 36.7 (2025): e70183.

13. Neri, M., Benevento, E., Stefanini, A., Aloini, D., Niccolini, F., Carducci, A., & Dini, G. "Understanding information security awareness: evidence from the public healthcare sector." *Information & Computer Security* 33.3 (2025): 309-319.
14. Nnadi, K., & Opoku, J. A. "Cybersecurity curriculum alignment with industry needs: A literature review of educational models integrating labs, certifications, and research." *Journal Of Engineering And Computer Sciences* 4.12 (2025): 64-76.
15. Pavlova, P., & Albert, C. D. "Defending Health Security." *The Cyber Defense Review* 10.2 (2025): 41-68.
16. Qureshi, R., & Koo, I. "A Comprehensive Survey of Cybersecurity Threats and Data Privacy Issues in Healthcare Systems." *Applied Sciences* 16.3 (2026): 1511.
17. Roumani, Y., & Roumani, Y. F. "Predicting Ransomware Incidents with Time-Series Modeling." *Journal of Cybersecurity and Privacy* 5.3 (2025): 61.
18. Salama, R., Altrjman, C., & Al-Turjman, F. "Healthcare cybersecurity challenges: a look at current and future trends." *Computational intelligence and Blockchain in complex systems* (2024): 97-111.

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Alornyo, S. Y. "Hospital Cyber Attack Forecasting: A Review of Time-Series Methods Used to Predict Security Incidents" *Sarcouncil Journal of Engineering and Computer Sciences* 5.4 (2026): pp 118-130.

Publisher: SARC Publisher