International Journal of Accounting, Finance and Risk Management
2026, Vol. 11, No. 3, pp. 138–147
https://doi.org/10.11648/j.ijafrm.20261103.13

SciencePG
Science Publishing Group

Research Article

Quantifying Cyber Risk Exposure and Risk-Based Pricing of Cyber Insurance: A Thematic Review of Actuarial, Statistical and Machine-Learning Approaches

Raveendran Narasimhan*

General Insurance, National Insurance Academy, Pune, India

Abstract

Cyber risk has evolved from a niche operational hazard into one of the most consequential and least tractable classes of insurable risk. Insurers continue to rely largely on heuristic and experience-rated underwriting, while the academic literature has produced a rapidly expanding but fragmented body of quantitative models. This review synthesises the peer-reviewed literature on cyber risk quantification and cyber insurance pricing, drawing on a structured search of scholarly databases that returned 199 records, from which 102 studies are critically reviewed. The literature is organised into eight themes: the statistical properties of cyber losses; actuarial and mathematical pricing models; dependence modelling and systemic risk; machine-learning approaches to incident prediction and underwriting; technical risk-assessment frameworks and control-based premium adjustment; the economics of cyber insurance markets; sectoral heterogeneity in cyber exposure; and methods for modelling under data scarcity. Across themes, four persistent findings emerge: cyber losses are heavy-tailed and dynamically non-stationary; dependence among losses undermines classical portfolio diversification and is systematically under-modelled in practice; technical vulnerability metrics and financial loss models remain poorly integrated; and virtually all empirical evidence derives from United States loss databases, leaving emerging markets and firm-level primary data almost unexamined. These gaps motivate an integrated, industry-specific research agenda that couples asset-centric technical assessment (CVSS, FAIR, ISO/NIST alignment) with neural-network incident probability estimation, copula-based dependence modelling and explicit actuarial premium construction. The review concludes by positioning such a framework against the state of the art and identifying the contribution it would make to actuarial science, underwriting practice and regulation.

Keywords

Cyber Risk, Cyber Insurance, Risk-based Pricing, Heavy Tails, Copula Models, Neural Networks, Systemic Risk

1. Introduction

The rapid digitalisation of economic activity, accelerated by cloud computing, application programming interfaces, third-party service networks, the Internet of Things (IoT) and, most recently, agentic artificial intelligence, has expanded the attack surface of organisations faster than the capacity of insurers to measure it. The economic significance of the resulting exposure is no longer disputed: empirical studies consistently document material market-value losses following cyber incidents [20, 42],

*Correspondence: Raveendran Narasimhan (raveendran@xsentinel.co)

Received: 19 July 2026; **Accepted:** 26 August 2026; **Published:** 18 September 2026

regulators increasingly treat cyber events as a source of systemic financial risk, and estimated economy-wide losses exceed the total capacity of the global cyber insurance market by an order of magnitude [15, 35] Yet the market designed to absorb this risk remains small, concentrated and, by most academic assessments, inefficiently priced [35, 60, 84].

The central difficulty is well characterised in the literature: cyber risk violates, in whole or in part, most of the classical insurability criteria. Losses are heavy-tailed, dependence among policyholders is pervasive, the threat environment is non-stationary, historical data are scarce, biased and rapidly obsolete, and severe information asymmetries separate insurers from insureds [6, 11, 102]. In response, scholars have developed a wide range of quantitative approaches, from copula-based actuarial models [45] and epidemic network models [38, 100] to Hawkes processes [46, 75], extreme value methods [53, 92] and, increasingly, machine-learning and hybrid frameworks [29, 39, 67]. Parallel literatures in information security have produced technical scoring systems, notably the Common Vulnerability Scoring System (CVSS) [61] and the Factor Analysis of Information Risk (FAIR) framework [93], which quantify vulnerabilities and translate them into loss estimates, while the economics of information security literature has examined moral hazard, adverse selection and the incentive effects of insurance [30, 55, 79].

These literatures have, however, developed largely in isolation from one another. Actuarial models are rarely informed by firm-level technical vulnerability data; technical frameworks rarely produce outputs in a form usable for premium calculation; machine-learning studies are dominated by intrusion detection rather than financial loss estimation; and nearly all empirical calibration rests on a small number of United States breach databases. The purpose of this review is to synthesise these strands critically, to establish what is settled, what is contested and what is absent, and thereby to motivate a research agenda for an integrated, industry-specific cyber risk quantification and risk-based pricing framework that combines asset-centric technical assessment, probabilistic and machine-learning modelling, dependence structures and explicit premium construction. The remainder of the paper proceeds as follows. Section 2 describes the review method. Section 3 reviews the literature thematically. Section 4 synthesises the findings and articulates the research gap, and Section 5 concludes.

2. Review Scope and Method

The review followed a structured, theme-driven search strategy executed through the Consensus academic search platform, which indexes more than 200 million peer-reviewed records across Semantic Scholar, Scopus, PubMed and arXiv. Ten themed queries were run covering: (i) actuarial pricing of cyber insurance; (ii) frequency-severity and heavy-tailed loss modelling; (iii) copula and dependence modelling; (iv) machine-learning prediction of cyber incidents; (v) cyber insurance market structure, demand and insurability; (vi) moral hazard, security investment and insurance economics; (vii) technical risk-assessment frameworks (NIST (National Institute of Standards and Technology), FAIR,(Factor Analysis of Information Risk) CVSS (common Vulnerability Scoring system), OCTAVE (Operationally critical threat, Asset and Vulnerability Evaluation); (viii) sector-specific exposure in healthcare, critical infrastructure and industrial control systems; (ix) systemic risk, accumulation, reinsurance and alternative risk transfer; and (x) expert judgment and Bayesian methods under data scarcity. The searches returned 199 records, which were de-duplicated and screened on title and abstract for direct relevance to cyber risk quantification or cyber insurance pricing.

Studies were retained if they (a) proposed or empirically evaluated a model for cyber loss frequency, severity, dependence or pricing; (b) provided empirical evidence on the statistical properties or economic cost of cyber incidents; (c) analysed the functioning of cyber insurance markets; or (d) developed technical or expert-judgment frameworks with demonstrable application to financially oriented cyber risk assessment. Purely technical intrusion-detection studies without a risk-quantification or insurance dimension were excluded, as were non-scholarly industry reports. One hundred and two studies satisfied these criteria and form the corpus of this review. The corpus spans 2003 to 2026, with a pronounced concentration after 2018, reflecting the recency of the field. Given the heterogeneity of methods and outcomes, a narrative thematic synthesis was adopted rather than a quantitative meta-analysis; themes were derived inductively from the corpus and correspond to Sections 3.1 to 3.8 below.

3. Thematic Review of the Literature

3.1. The Statistical Properties of Cyber Losses

3.1.1. Heavy Tails and Extreme Losses

The single most robust empirical finding in the corpus is that cyber loss severity is heavy-tailed. Maillart and Sornette [59] provided the earliest systematic evidence, documenting a remarkably stable power-law tail for personal identity losses per event, with a tail exponent below one, implying that aggregate losses are dominated by the largest events. Subsequent work has confirmed heavy-tailedness across datasets, sectors and periods. Edwards et al. [32], using Bayesian generalised linear models on public breach data, found breach sizes to be log-normally distributed with a negative binomial frequency, and showed that apparent growth in breach severity is largely an artefact of the heavy tail rather than a worsening trend. Wheatley et al. [97], analysing United States breach data within a man-made catastrophe framework, reached the opposite conclusion for externally caused hack events, finding both frequency and severity of large breaches worsening; the tension between these findings illustrates the sensitivity of tail inference to dataset, period and event definition. Analyses of

operational-risk style loss data consistently favour generalized Pareto or similarly heavy-tailed severity models for the tail [15, 87, 89], while Dacorogna et al. [28], using French national gendarmerie complaint data and a novel algorithm for asymmetric heavy-tailed data, confirmed finite loss expectation, a necessary condition for insurability. Jung [53] applied generalized extreme value methods to maxima of data breach losses and found the probable maximum loss to be substantially larger than Pareto-based estimates in the earlier literature, with a structural break after 2014, and von Skarczinski et al. [92] developed a tempered GEV approach on a stratified survey of 5,000 German organisations, one of the few studies to escape dependence on public United States databases.

3.1.2. Frequency, Severity and Their Interaction

At the firm level, the canonical frequency-severity decomposition has been progressively refined. Sun et al. [89] showed that hacking breach frequency at company level is better described by a hurdle Poisson model than the negative binomial commonly assumed, and documented a positive non-linear dependence between frequency and severity, contradicting the independence assumption embedded in standard collective risk models. Lu et al. [58] extended this line to a discrete multivariate count framework accommodating zero inflation, serial and cross-sectional correlation and heavy tails. Eling and Jung [34] demonstrated with quantile regressions that covariate effects differ sharply across the loss distribution: firm revenue matters more in the lower quantiles and less in the upper quantiles, so that pricing small and large firms from mean-effect models systematically misprices both. Shevchenko et al. [80, 81], analysing 132,126 events in the Advisen database, found that frequency and severity depend jointly on business sector and threat type, that more than 60 per cent of loss-recording firms suffered repeat events, and, importantly, that there is no clear proportional relationship between total loss and the number of compromised records, undermining record-count-based exposure proxies that remain common in practice.

3.1.3. Non-stationarity and Data Biases

Cyber loss processes are dynamically unstable in ways that natural-catastrophe processes are not. Eling et al. [36] documented exponentially growing frequency of malicious events over two decades with persistent heavy-tailedness but no significant change in severity, and used an extended two-stage model to correct for reporting delay, a bias that inflates apparent recent-period declines. Eling et al. [37] provide the most recent structural-change analysis of severity and tail dynamics. The practical implication, emphasised across the corpus, is that unadjusted historical calibration systematically understates current exposure, and that any pricing framework must incorporate forward-looking components, whether through expert judgment (Section 3.8), dynamic models (Section 3.2) or frequently recalibrated machine-learning components (Section 3.4).

3.2. Actuarial and Mathematical Models for Cyber Insurance Pricing

The actuarial literature has moved through three recognisable generations. The first generation adapted classical tools directly. Herath and Herath [45] developed the first copula-based actuarial model for cyber insurance, pricing first-party virus losses by coupling standard insurance risk elements with a copula dependence structure, and Bohme [13] established early on that platform monoculture induces correlated losses requiring premium surcharges. The second generation introduced explicitly dynamic and network-aware structures. Fahrenwaldt et al. [38] modelled infectious cyber threats spreading over a data network as an interacting Markov chain with claims as a marked point process, showing that network topology materially affects aggregate losses and that higher-order approximations are indispensable for non-linear claims. Xu and Hua [100] combined epidemic models, loss functions and premium principles into a simulation-based pricing framework, and Antonio et al. [3, 4] extended Markov-based epidemic pricing to incorporate network clustering structure and weighted graph-mining characteristics, producing heterogeneous, competitively priced premiums. Zeller and Scherer [102] proposed what is arguably the most complete second-generation model: a marked point process framework distinguishing untargeted attacks, targeted attacks and accidental failures, with covariate-driven frequency and severity and explicit treatment of accumulation.

The third generation addresses clustering in time and model risk. Hillairet et al. [46] derived an expansion formula for Hawkes processes enabling premium bounds for cyber reinsurance contracts under self-exciting attack arrivals, and Ren et al. [75] refined the Hawkes framework by allowing loss magnitude to feed back into the conditional intensity. Peters et al. [71] posed a question of direct practical significance: does model risk exist in cyber loss data and how does it translate into premium mispricing? Using robust estimators for marginal and joint loss processes they showed that parameter and model uncertainty produce economically material mispricing, a finding that should temper confidence in any single parametric specification. Complementing model-based work, Woods et al. [98] inverted the problem by extracting loss distributions from 6,828 observed premiums in United States regulatory filings, deriving a market-implied 'County Fair' loss distribution, and Bardopoulos [8] provided a comprehensive actuarial treatment of nonparametric severity models with large-loss components and increased limit factors under correlated aggregate loss. Recent contributions engage portfolio-level constraints directly: Jeamaon et al. [52] optimised premiums under Value-at-Risk and Tail Value-at-Risk constraints with a spliced lognormal-GPD severity and systemic common-shock dependence, finding that systemic events raise TVaR by 75.6 per cent while raising expected loss by under 6 per cent, so that risk-adequate premiums sit roughly 150 per cent above expected loss. Chong et al. [23] built the economic

foundation for incident-specific bundled coverage with per-peril limits and deductibles, and application-oriented studies have extended pricing to particular infrastructures and markets, including IP-based power substations using ruin-theoretic premium pools [101], security-maturity-model-based pricing for organisational systems [85] and Monte Carlo scenario pricing calibrated to the Egyptian market [50]. Comprehensive surveys of this modelling landscape are provided by Awiszus et al. [6], who distinguish idiosyncratic, systematic and systemic cyber risk and show that classical actuarial valuation suffices only for the first two, and by Carannante et al. [19], who review vulnerability functions through to stochastic dynamic models and reinsurance structuring.

Taken together, this literature is technically mature but operationally incomplete. The models are calibrated almost exclusively to public breach databases rather than to underwriting-relevant firm characteristics; covariates describing an insured's actual control environment enter, if at all, through coarse proxies. The bridge between what an underwriter can observe about a specific firm and the parameters of these models remains the weakest link, a gap taken up in Sections 3.4 and 3.5.

3.3. Dependence, Accumulation and Systemic Risk

Because cyber events propagate through shared software, shared service providers and interconnected networks, dependence is the defining actuarial feature of the risk class. Eling and Jung [33] provided the first substantial empirical evidence, identifying significant asymmetric cross-sectional dependence of monthly data breach losses across industries and breach types using pair copula constructions. Peng et al. [70] showed with a Copula-GARCH vine model of attack data that ignoring multivariate dependence severely underestimates cyber risk, and Wu et al. [99] coupled deep learning point prediction with extreme value theory for high-quantile prediction of multivariate cyber risks. At portfolio level, Carannante et al. [18] used vine copulas to expose a regulatory paradox under Solvency II, whereby an insurer can appear solvent under Pillar 1 while materially under-capitalised for operational cyber dependence under Pillar 2.

A second stream models accumulation mechanisms structurally rather than statistically. Hillairet and Lopez [47] combined counting processes with compartmental epidemiological models to design accumulation scenarios mimicking a WannaCry-type event across an insurance portfolio, quantifying the value of response countermeasures. Welburn and Strong [96] extended sector-level input-output analysis to cascading cyber failures, estimating aggregate economic damages from firm-level incidents. Zeller and Scherer [103] delivered a caution of particular importance for practice: when claims data collection is not aligned with the dependence model, naive data collection necessarily leads to underestimation of accumulation risk, because common-vulnerability dependence is invisible in loss records that do not tag root causes. Pal et al. [68, 69] established, via game-theoretic and statistical analysis of heavy-tailed, tail-dependent risks in IoT-driven supply chains, conditions under which aggregate cyber-risk coverage and reinsurance are economically infeasible without government intervention. The risk-transfer response is examined by Braun et al. [16], who assessed the feasibility of cyber insurance-linked securities and found that a meaningful market requires modelling maturity that does not yet exist, and by Cremer et al. [26], whose interviews with twenty reinsurance market participants document that 50 to 65 per cent of global cyber premiums were ceded to reinsurers in 2022 under stringent capacity criteria. The consistent conclusion across this theme is that dependence modelling is not a refinement but a first-order determinant of capital, pricing and insurability; frameworks that treat firm-level exposures as independent will misstate portfolio risk by construction.

3.4. Machine Learning and Data-Driven Approaches

Machine learning enters the literature along three distinct paths. The first is incident and rate prediction. Fang et al. [39] developed a bidirectional LSTM framework for forecasting cyber attack rates, exploiting long-range dependence and nonlinearity that defeat classical time-series models, and Samia et al. [78] combined real-time data collection with recurrent and convolutional architectures to forecast attack frequency within time windows, reporting accuracy gains exceeding 15 per cent over baselines. Studies of this type demonstrate that attack arrival processes are learnable to a useful degree, which matters for pricing because frequency parameters need not be static. Bilen and Ozer [12] further showed that attack methods and perpetrator characteristics are classifiable from incident features, though such forensic classification sits at some distance from underwriting.

The second path applies machine learning to underwriting and premium estimation in general insurance, where tree ensembles and neural networks now routinely outperform generalised linear benchmarks in risk classification, with SHAP-based interpretability increasingly used to satisfy regulatory explainability requirements [76]. The third and most relevant path is the emerging hybrid literature that couples machine learning with established risk-quantification frameworks. Nwafor et al. [67] integrated the FAIR taxonomy with XGBoost and SHAP explainability, constructing a composite risk exposure score from frequency, vulnerability, control maturity, severity and downtime, with NIST CSF-based control maturity quantified through weighted technology, process and people scores; expected annual loss was found to grow exponentially beyond a threshold exposure score. Guo et al. [43] addressed the covariate gap directly by enriching incident data with entity-specific organisational features from InsurTech

sources, showing that such features materially improve occurrence and frequency estimation across incident types. Most directly aligned with the agenda of this review, Doss et al. [29] developed a multilayer neural network estimating the probability of loss for cyber risks across industry segments, combined with structural equation modelling of aggregate exposure effects, finding that system vulnerability, user awareness and mitigation practices significantly determine exposure and that model outputs support actuarial pricing by sector. Dusane et al. [31], reviewing ten academic and industry models, concluded that the binding constraints on current practice are non-standardised incident data, the limitations of traditional actuarial techniques and the failure to integrate technical, human and organisational factors, for which adaptive underwriting, Bayesian modelling and machine-learned prediction are the emerging remedies. The critical reading of this theme is that machine learning contributes most where it is disciplined by an economic or actuarial structure; unstructured classifier accuracy on benchmark intrusion datasets translates poorly into premium-relevant quantities, and interpretability constraints bind hard in regulated insurance applications.

3.5. Technical Risk-Assessment Frameworks and Control-Based Premium Adjustment

If actuarial models supply the pricing machinery, technical frameworks supply the firm-specific inputs. CVSS remains the de facto standard for vulnerability severity measurement [61], and a substantial engineering literature builds quantitative risk-assessment methodologies on it, incorporating asset centrality, attack graphs and exploitation frequency [2, 54]. The framework's limitations are, however, well documented from within the security community: Spring et al. [86] demonstrated that the CVSS scoring algorithm is neither formally nor empirically justified and warned explicitly against its direct use as a risk score, and Howland [48] found no correlation between CVSS scores and vulnerabilities actually exploited in the wild. For an insurance application, the implication is not that CVSS should be discarded, since it is standardised, universally available and auditable, but that it must enter as one covariate among several rather than as a risk measure in itself.

FAIR occupies the complementary role of translating technical scenarios into financial loss distributions through structured decomposition of loss event frequency and magnitude. Wang et al. [93] implemented and extended FAIR with Bayesian networks, removing the distributional restrictions and approximation error of the native Monte Carlo implementation and demonstrating extensibility to process-oriented and game-theoretic components. Bayesian network approaches more broadly constitute the most developed formalism for integrating technical architecture with probabilistic risk: applications span nuclear instrumentation and control [82], dynamic assessment in industrial control systems with fuzzy probabilities compensating for scarce data [104], and decision-analytic frameworks that integrate threat, vulnerability and consequence into management-ready metrics [40]. Governance-oriented instruments complete the input set: NIST CSF-based self-assessment tools have been operationalised for small and medium enterprises [10], and IoT-specific economic impact models map Industry 4.0 technology trends into insurable exposure categories [74]. What no published study yet accomplishes is the systematic fusion of these instruments into a premium formula: CVSS scores, FAIR loss decomposition, ISO/NIST governance maturity and asset valuation are each used somewhere, but no framework carries all of them jointly through to an explicit, auditable premium with a control-based adjustment term. This integration gap is the pivotal opening for new research.

3.6. The Economics of Cyber Insurance: Insurability, Information and Incentives

The economics literature explains why the cyber insurance market remains small relative to exposure. Biener et al. [11], applying Berliner's insurability criteria to 994 cyber loss events, identified interrelated losses, data scarcity and information asymmetry as the binding constraints. Eling et al. [35] showed that because cyber risk is simultaneously information-intensive to underwrite and heavy-tailed, supply concentrates in large insurance groups with deep internal capital markets, and established causally that internal capital supports cyber capacity. Skeoch and Bohme [84] demonstrated through market simulation that absent data sharing, limited reinsurer involvement raises premiums and suppresses capacity. On the demand side, Cremer et al. [27] documented pervasive non-standardisation of policy wordings across 41 policies covering roughly 80 per cent of the German market; Salzberger [77] found in a survey of 1,248 German SMEs that anxiety and estimated financial impact drive purchase while probability perceptions do not, evidencing the difficulty firms have in estimating incident probabilities; and Ning [66] documented two-stage adverse selection in the United States broker market, with insurers responding by capping coverage quantities rather than differentiating prices, precisely because their survey-based screening correlates only weakly with underlying risk. Emerging-market evidence remains thin: Koley [56] surveys the Indian market, identifying awareness gaps, pricing complexity and non-standardised products as the principal barriers, and Subramaniam et al. [88] rank market drivers and barriers using fuzzy multi-criteria methods, but neither engages primary firm-level exposure data.

The incentive literature yields a nuanced verdict on whether insurance improves security. Early network-economics models were optimistic that insurance would function as an incentive mechanism for self-protection [14, 57], but equilibrium analyses show that with unobservable security effort, moral hazard erodes or eliminates the market and can worsen network security [79]. The corrective mechanisms studied since are directly relevant to pricing design: security pre-screening

with premium discrimination can restore incentives and improve security beyond the no-insurance baseline [55]; optimal contracts can be built to maximise insured utility under ex ante moral hazard [30]; and integrated models of security investment and insurance, in the Gordon-Loeb tradition, identify when the two are complements rather than substitutes [83, 94]. Empirically, the ransomware experience is cautionary: Baker and Shortland [7] found that insurers developed sophisticated ex post loss-reduction services while leaving ex ante security decisions to insureds, thereby facilitating enterprise but arguably fuelling the ransom economy, and Mott et al. [62], drawing on 96 practitioner interviews, describe insurers caught between risk-based pricing pressure and competitive underwriting laxity. Mukhopadhyay et al. [64] respond with a ransomware-specific risk management model combining predictive risk assessment, collective risk quantification and residual-risk transfer. The through-line for pricing research is that a credible control-adjustment term in the premium is not merely an actuarial refinement; it is the principal instrument by which insurance can address moral hazard and function as private governance of cyber security.

3.7. Sectoral Heterogeneity of Cyber Exposure

The evidence is unambiguous that cyber exposure is industry-specific in frequency, severity, threat mix and consequence, which is the empirical premise of industry-specific pricing. In the cross-sector loss databases, the information, professional services and financial sectors absorb most recorded financial damage, while threat-type composition varies sharply by sector [80, 81]. Event studies add a market-value dimension: early studies documented significant negative abnormal returns concentrated in breaches of confidential data [1, 17, 20, 42], with subsequent work showing that firm size and disclosure behaviour moderate losses [41, 90]. Celeny et al. [21], re-estimating with corrected standard errors over 2012-2022, found data breaches to be the most damaging incident type and the health sector the most price-sensitive, with average abnormal losses exceeding 5 per cent, and Muktadir-Al-Mukit et al. [63] estimate an average market-value loss of 309 million dollars on announcement day, amplified by media salience, severity and first-time occurrence.

Sector studies deepen the picture. Healthcare exhibits a distinctive combination of highly monetisable data, safety-critical operations, device-level attack surfaces and chronic underinvestment [5, 25, 51, 95]; externally observable security ratings predict hospital breach probability, with annual breach probabilities of 14 to 33 per cent for low-rated hospitals [22], and national-health-system case studies quantify direct operational costs of single incidents in the millions of euros [73]. Manufacturing, energy and utilities face a different profile dominated by industrial control system and SCADA vulnerabilities, where cyber events cause physical disruption and business interruption rather than data liability [65, 104], and IoT proliferation couples ransomware to operational technology [49]. Financial institutions face the heaviest regulatory and third-party liability exposure, with operational-risk-based estimates of sectoral cyber losses far exceeding market premium volumes [15]. This heterogeneity has a direct modelling consequence documented repeatedly in the statistical literature: pooled cross-sector calibration biases both frequency and severity parameters, yet most published pricing models remain sector-agnostic, treating industry at best as a categorical covariate rather than modelling sector-specific threat vulnerability structures.

3.8. Modelling Under Data Scarcity: Expert Judgment and Bayesian Methods

Because credible firm-level loss histories rarely exist, especially outside the United States, methods for disciplined incorporation of expert knowledge are integral to cyber risk quantification rather than a second-best. The Bayesian network literature provides the canonical machinery: Podofillini et al. [72] set out a traceable process for building networks from scarce data and expert judgment with sequential Bayesian updating; Constantinou et al. [24] solved the problem of adding expert variables to data-driven networks while preserving empirically observed distributions; and Barons et al. [9] evaluated interpolation methods that reduce the elicitation burden when conditional probability tables are large, a practical constraint in any realistic cyber application. Hanea et al. [44] survey the broader case for Bayesian networks in risk analysis and decision support under uncertainty. Applied demonstrations in adjacent safety-critical domains show the approach scales to operational use, notably Uflaz et al. [91], who combined Dempster-Shafer theory with FMECA and rule-based Bayesian networks to quantify maritime cyber attack risks from structured expert evaluations. Within insurance-oriented cyber modelling, fuzzy-probability variants address the same problem in industrial control settings [104], and the FAIR-BN integration of Wang et al. [93] shows how elicited quantities propagate coherently into loss distributions. Structured elicitation methods of the Delphi type are widely invoked in applied cyber risk studies but are rarely reported with methodological rigour; the transfer of formal structured expert judgment protocols from risk analysis into cyber insurance calibration remains visibly underdeveloped, and constitutes a methodological opportunity for studies that combine practitioner panels, such as chief information security officers and underwriters, with quantitative modelling.

4. Synthesis and Research Gap

Four conclusions can be stated with confidence from the 102 studies reviewed. First, the statistical character of cyber loss is settled in outline: severity is heavy-tailed with finite mean, frequency is rising and non-stationary, frequency and severity are dependent, and covariate effects vary across the

loss distribution. Second, the actuarial toolkit is rich but calibrated to thin, biased, predominantly American public data; model risk translates directly into premium mispricing. Third, dependence and accumulation are the decisive portfolio-level phenomena, and their under-measurement is structural, arising from data collection practices that do not record common causes. Fourth, market frictions, information asymmetry, moral hazard, adverse selection and non-standardised products are as binding as statistical difficulties, and premium design that rewards verified controls is the principal instrument available to relax them.

Against these findings, five gaps jointly define the opening for the research programme this review supports. (i) Integration: no published framework carries asset-centric technical assessment (CVSS-scored vulnerabilities, OCTAVE-style asset criticality, ISO 27014 governance maturity), FAIR-based financial translation, machine-learned incident probability and copula-based dependence jointly through to an explicit premium decomposition of the form expected loss plus risk loading plus or minus control adjustment plus expenses and margin. The components exist separately [29, 52, 67, 93]; their end-to-end fusion does not. (ii) Primary data: virtually all empirical work relies on secondary breach databases; survey-based primary evidence from security officers exists only in isolated national studies [77, 92], and none links survey-measured technical posture to loss modelling at the industry level. (iii) Sector specificity: the documented heterogeneity of threat profiles across banking and financial services, IT services, telecommunications, manufacturing and energy, and healthcare has not been carried into sector-specific pricing structures; industry enters existing models as a dummy variable rather than as a distinct vulnerability architecture. (iv) Emerging markets: the Indian market, and emerging markets generally, are described qualitatively [56, 88] but wholly unmodelled quantitatively, despite regulatory developments such as the Digital Personal Data Protection Act materially altering the liability structure that premiums must reflect. (v) Incentive-compatible pricing: the theoretical case for control-contingent premiums is established [30, 55], but no empirical framework operationalises a control adjustment factor from auditable firm-level measurements. A study that addresses these five gaps simultaneously, quantifying industry-specific exposure from primary CISO-level data, estimating incident probabilities with neural networks validated against expert elicitation, modelling dependence with copulas, and expressing the result as a transparent risk-based premium, would occupy ground that the existing literature demonstrably has not.

5. Conclusion

The academic study of cyber risk quantification and insurance pricing has matured rapidly, producing robust stylised facts about loss behaviour, a sophisticated repertoire of actuarial and machine-learning models, and a clear economic diagnosis of the market's frictions. What it has not yet produced is an integrated, empirically grounded, industry-specific framework that connects what can be measured inside a firm to what an insurer must charge, in a form that is transparent to regulators and incentive-compatible for insureds. The literature reviewed here defines both the necessity of such a framework and the components from which it can be assembled. Closing that gap is the object of the research programme to which this review is a prelude: the quantification of cyber risk exposures in selected industries and the development of a risk-based pricing model, with implications for customers, insurers and regulators alike.

Abbreviations

NIST	National Institute of Standards and Technology
FAIR	Factor Analysis of Information Risk
CVSS	Common Vulnarability Scoring System
SCADA	Supervisory Control and Data Acquisition
OCTAVE	Operationally Critical Threat, Asset and Vulnerability Evaluation

Author Contributions

Raveendran Narasimhan: Conceptualization, Writing – original draft

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Acquisti, A., Friedman, A. and Telang, R. (2006). Is there a cost to privacy breaches? An event study. Proceedings of the International Conference on Information Systems / Workshop on the Economics of Information Security.

[2] Aksu, M. U., et al. (2017). A quantitative CVSS-based cyber security risk assessment methodology for IT systems. Proceedings of the International Carnahan Conference on Security Technology (ICCST).

[3] Antonio, Y., et al. (2021a). Pricing of cyber insurance premiums using a Markov-based dynamic model with clustering structure. PLoS ONE, 16.

[4] Antonio, Y., et al. (2021b). Cyber insurance ratemaking: A graph mining approach. Risks, 9.

[5] Argaw, S. T., et al. (2020). Cybersecurity of hospitals: Discussing the challenges and working towards mitigating the risks. BMC Medical Informatics and Decision Making, 20.

[6] Awiszus, K., et al. (2023). Modeling and pricing cyber insurance: Idiosyncratic, systematic, and systemic risks. European Actuarial Journal, 13.

[7] Baker, T. and Shortland, A. (2022). Insurance and enterprise: Cyber insurance for ransomware. The Geneva Papers on Risk and Insurance - Issues and Practice, 48.

[8] Bardopoulos, J. (2025). Cyber-insurance pricing models. British Actuarial Journal, 30.

[9] Barons, M. J., et al. (2021). Balancing the elicitation burden and the richness of expert input when quantifying discrete Bayesian networks. Risk Analysis, 42.

[10] Benz, M. and Chatterjee, D. (2020). Calculated risk? A cybersecurity evaluation tool for SMEs. Business Horizons, 63.

[11] Biener, C., Eling, M. and Wirfs, J. H. (2014). Insurability of cyber risk: An empirical analysis. The Geneva Papers on Risk and Insurance - Issues and Practice, 40.

[12] Bilen, A. and Ozer, A. B. (2021). Cyber-attack method and perpetrator prediction using machine learning algorithms. PeerJ Computer Science, 7.

[13] Bohme, R. (2005). Cyber-insurance revisited. Proceedings of the Workshop on the Economics of Information Security (WEIS).

[14] Bolot, J. and Lelarge, M. (2008). Cyber insurance as an incentive for IT security. Proceedings of the Workshop on the Economics of Information Security (WEIS).

[15] Bouveret, A. (2019). Estimation of losses due to cyber risk for financial institutions. Journal of Operational Risk, 14.

[16] Braun, A., et al. (2023). Cyber insurance-linked securities. ASTIN Bulletin, 53.

[17] Campbell, K., et al. (2003). The economic cost of publicly announced information security breaches: Empirical evidence from the stock market. Journal of Computer Security, 11.

[18] Carannante, M., et al. (2023). Vine copula modelling dependence among cyber risks: A dangerous regulatory paradox. SSRN Working Paper.

[19] Carannante, M., et al. (2025). An analytical review of cyber risk management by insurance companies: A mathematical perspective. Risks, 13.

[20] Cavusoglu, H., Mishra, B. and Raghunathan, S. (2004). The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. International Journal of Electronic Commerce, 9.

[21] Celeny, D., et al. (2024). Prioritizing investments in cybersecurity: Empirical evidence from an event study on the determinants of cyberattack costs. SSRN Working Paper.

[22] Choi, S. J., et al. (2021). The relationship between cybersecurity ratings and the risk of hospital data breaches. Journal of the American Medical Informatics Association, 28.

[23] Chong, W. F., et al. (2023). Incident-specific cyber insurance. ASTIN Bulletin, 53.

[24] Constantinou, A., Fenton, N. and Neil, M. (2016). Integrating expert knowledge with data in Bayesian networks: Preserving data-driven expectations when the expert variables remain unobserved. Expert Systems with Applications, 56.

[25] Coventry, L. and Branley, D. (2018). Cybersecurity in healthcare: A narrative review of trends, threats and ways forward. Maturitas, 113.

[26] Cremer, F., et al. (2024a). Enhancing cyber insurance strategies: Exploring reinsurance and alternative risk transfer approaches. Journal of Cybersecurity, 10.

[27] Cremer, F., et al. (2024b). Bridging the cyber protection gap: An investigation into the efficacy of the German cyber insurance market. Risk Management and Insurance Review, 27.

[28] Dacorogna, M., et al. (2022). Building up cyber resilience by better grasping cyber risk via a new algorithm for modelling heavy-tailed data. European Journal of Operational Research, 311.

[29] Doss, S., et al. (2026). Development of an aggregate model for cyber risk assessment using deep neural network and structural equation modelling. International Journal of Finance and Economics.

[30] Dou, W., et al. (2020). An insurance theory based optimal cyber-insurance contract against moral hazard. Information Sciences, 527.

[31] Dusane, H. N., et al. (2025). Evaluating cyber risk insurance frameworks: Bridging cybersecurity threats with financial risk strategies through actuarial modeling and adaptive resilience. Proceedings of the IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS).

[32] Edwards, B., Hofmeyr, S. and Forrest, S. (2016). Hype and heavy tails: A closer look at data breaches. Journal of Cybersecurity, 2.

[33] Eling, M. and Jung, K. (2018). Copula approaches for modeling cross-sectional dependence of data breach losses. Insurance: Mathematics and Economics, 82.

[34] Eling, M. and Jung, K. (2022). Unraveling heterogeneity in cyber risks using quantile regressions. Insurance: Mathematics and Economics, 104.

[35] Eling, M., et al. (2023a). The supply of cyber risk insurance. SSRN Working Paper.

[36] Eling, M., et al. (2023b). Time dynamics of cyber risk. SSRN Working Paper.

[37] Eling, M., et al. (2025). The changing landscape of cyber risk: An empirical analysis of loss severity and tail dynamics. Insurance: Mathematics and Economics.

[38] Fahrenwaldt, M., Weber, S. and Weske, K. (2018). Pricing of cyber insurance contracts in a network model. ASTIN Bulletin, 48.

[39] Fang, X., et al. (2019). A deep learning framework for predicting cyber attacks rates. EURASIP Journal on Information Security, 2019.

[40] Ganin, A., et al. (2020). Multicriteria decision framework for cybersecurity risk assessment and management. Risk Analysis, 40.

[41] Gatzlaff, K. M. and McCullough, K. A. (2010). The effect of data breaches on shareholder wealth. Risk Management and Insurance Review, 13.

[42] Goel, S. and Shawky, H. A. (2009). Estimating the market impact of security breach announcements on firm values. Information and Management, 46.

[43] Guo, J., et al. (2025). Entity-specific cyber risk assessment using InsurTech empowered risk factors. arXiv preprint.

[44] Hanea, A., et al. (2022). Bayesian networks for risk analysis and decision support. Risk Analysis, 42.

[45] Herath, H. S. B. and Herath, T. C. (2011). Copula-based actuarial model for pricing cyber-insurance policies. Insurance Markets and Companies: Analyses and Actuarial Computations, 2.

[46] Hillairet, C., et al. (2021a). An expansion formula for Hawkes processes and application to cyber-insurance derivatives. Stochastic Processes and their Applications, 143.

[47] Hillairet, C. and Lopez, O. (2021). Propagation of cyber incidents in an insurance portfolio: Counting processes combined with compartmental epidemiological models. Scandinavian Actuarial Journal, 2021.

[48] Howland, H. (2021). CVSS: Ubiquitous and broken. Digital Threats: Research and Practice, 4.

[49] Humayun, M., et al. (2020). Internet of things and ransomware: Evolution, mitigation and prevention. Egyptian Informatics Journal, 22.

[50] Ismail, S. M., et al. (2025). Cyber insurance pricing through Monte Carlo simulation: A case study of the Egyptian insurance market. Journal of Administrative, Financial and Quantitative Research.

[51] Jalali, M. S. and Kaiser, J. P. (2018). Cybersecurity in hospitals: A systematic, organizational perspective. Journal of Medical Internet Research, 20.

[52] Jeamaon, A., et al. (2026). Aggregate cyber loss modeling with TVaR-constrained premium optimization: A Monte Carlo framework with spliced severity and systemic dependence. Proceedings of the IEEE International Conference on Cybernetics and Innovations (ICCI).

[53] Jung, K. (2021). Extreme data breach losses: An alternative approach to estimating probable maximum loss for data breach risk. North American Actuarial Journal, 25.

[54] Keskin, O., et al. (2021). Scoring cyber vulnerabilities based on their impact on organizational goals. Proceedings of the Systems and Information Engineering Design Symposium (SIEDS).

[55] Khalili, M. M., Naghizadeh, P. and Liu, M. (2017). Designing cyber insurance policies: Mitigating moral hazard through security pre-screening. Proceedings of GameNets / International Conference on Game Theory for Networks.

[56] Koley, J. (2025). Emerging trends and challenges in India's cyber insurance sector: A multifaceted examination. International Journal of Financial Management and Economics, 8.

[57] Lelarge, M. and Bolot, J. (2009). Economic incentives to increase security in the internet: The case for insurance. Proceedings of IEEE INFOCOM 2009.

[58] Lu, Y., et al. (2024). Cyber risk modeling: A discrete multivariate count process approach. Scandinavian Actuarial Journal, 2024.

[59] Maillart, T. and Sornette, D. (2008). Heavy-tailed distribution of cyber-risks. The European Physical Journal B, 75.

[60] Marotta, A., et al. (2017). Cyber-insurance survey. Computer Science Review, 24.

[61] Mell, P., Scarfone, K. and Romanosky, S. (2006). Common Vulnerability Scoring System. IEEE Security and Privacy, 4.

[62] Mott, G., et al. (2023). Between a rock and a hard (ening) place: Cyber insurance in the ransomware era. Computers and Security, 128.

[63] Muktadir-Al-Mukit, D., et al. (2025). The dynamics of stock market responses following the cyber-attacks news: Evidence from event study. Information Systems Frontiers.

[64] Mukhopadhyay, A., et al. (2024). A framework for cyber-risk insurance against ransomware: A mixed-method approach. International Journal of Information Management, 74.

[65] Nankya, M., et al. (2023). Securing industrial control systems: Components, cyber threats, and machine learning-driven defense strategies. Sensors, 23.

[66] Ning, D. (2026). Selection and screening in cyber insurance markets. SSRN Working Paper.

[67] Nwafor, C., et al. (2025). A hybrid FAIR and XGBoost framework for cyber-risk intelligence and expected loss prediction. Expert Systems with Applications.

[68] Pal, R., et al. (2021a). Will catastrophic cyber-risk aggregation thrive in the IoT age? A cautionary economics tale for (re-) insurers and likes. ACM Transactions on Management Information Systems, 12.

[69] Pal, R., et al. (2021b). Aggregate cyber-risk management in the IoT age: Cautionary statistics for (re) insurers and likes. IEEE Internet of Things Journal, 8.

[70] Peng, C., et al. (2018). Modeling multivariate cybersecurity risks. Journal of Applied Statistics, 45.

[71] Peters, G. W., et al. (2022). Cyber loss model risk translates to premium mispricing and risk sensitivity. The Geneva Papers on Risk and Insurance - Issues and Practice, 48.

[72] Podofillini, L., et al. (2022). A traceable process to develop Bayesian networks from scarce data and expert judgment: A human reliability analysis application. Reliability Engineering and System Safety, 230.

[73] Portela, D., et al. (2022). Economic impact of a hospital cyberattack in a national health system: Descriptive case study. JMIR Formative Research, 7.

[74] Radanliev, P., et al. (2018). Economic impact of IoT cyber risk: Analysing past and present to predict the future developments in IoT risk analysis and IoT cyber insurance. arXiv preprint.

[75] Ren, N., et al. (2025). The modeling of cyber risk insurance by Hawkes processes with loss covariate. Applied Stochastic Models in Business and Industry, 41.

[76] Sahai, R., et al. (2023). Insurance risk prediction using machine learning. Proceedings of the International Conference on Data Science and Applications.

[77] Salzberger, A. (2025). An empirical analysis of the behavioral influences and information sources affecting the cyber insurance decisions of German SMEs. The Journal of Risk Finance, 26.

[78] Samia, N., et al. (2024). Predicting and mitigating cyber threats through data mining and machine learning. Computer Communications, 216.

[79] Shetty, N., et al. (2010). Competitive cyber-insurance and internet security. In: Moore, T., Pym, D. and Ioannidis, C. (eds.) Economics of Information Security and Privacy. Springer.

[80] Shevchenko, P. V., et al. (2021). Quantification of cyber risk: Risk categories and business sectors. SSRN Working Paper.

[81] Shevchenko, P. V., et al. (2022). The nature of losses from cyber-related events: Risk categories and business sectors. Journal of Cybersecurity, 9.

[82] Shin, J., Son, H. and Heo, G. (2015). Development of a cyber security risk model using Bayesian networks. Reliability Engineering and System Safety, 134.

[83] Skeoch, H. (2021). Expanding the Gordon-Loeb model to cyber-insurance. Computers and Security, 112.

[84] Skeoch, H. and Bohme, R. (2023). The barriers to sustainable risk transfer in the cyber-insurance market. Journal of Cybersecurity, 9.

[85] Skeoch, H., et al. (2023). Pricing cyber-insurance for systems via maturity models. arXiv preprint.

[86] Spring, J. M., et al. (2021). Time to change the CVSS? IEEE Security and Privacy, 19.

[87] Strupczewski, G. (2019). What is the worst scenario? Modeling extreme cyber losses. In: Multiple Perspectives in Risk and Risk Management. Springer Proceedings in Business and Economics.

[88] Subramaniam, B., et al. (2023). What ails cyber insurance? An analysis of barriers and drivers using fuzzy TOPSIS method. SN Computer Science, 4.

[89] Sun, H., Xu, M. and Zhao, P. (2020). Modeling malicious hacking data breach risks. North American Actuarial Journal, 25.

[90] Tripathi, M. and Mukhopadhyay, A. (2020). Financial loss due to a data privacy breach: An empirical analysis. Journal of Organizational Computing and Electronic Commerce, 30.

[91] Uflaz, E., et al. (2023). Quantifying potential cyber-attack risks in maritime transportation under Dempster-Shafer theory FMECA and rule-based Bayesian network modelling. Reliability Engineering and System Safety, 243.

[92] von Skarczinski, B. S., et al. (2023). Modelling maximum cyber incident losses of German organisations: An empirical study and modified extreme value distribution approach. The Geneva Papers on Risk and Insurance - Issues and Practice, 48.

[93] Wang, J., Neil, M. and Fenton, N. (2020). A Bayesian network approach for cybersecurity risk assessment implementing and extending the FAIR model. Computers and Security, 89.

[94] Wang, S. S. (2019). Integrated framework for information security investment and cyber insurance. Pacific-Basin Finance Journal, 57.

[95] Wasserman, L. and Wasserman, Y. (2022). Hospital cybersecurity risks and gaps: Review (for the non-cyber professional). Frontiers in Digital Health, 4.

[96] Welburn, J. W. and Strong, A. M. (2021). Systemic cyber risk and aggregate impacts. Risk Analysis, 42.

[97] Wheatley, S., Hofmann, A. and Sornette, D. (2020). Addressing insurance of data breach cyber risks in the catastrophe framework. The Geneva Papers on Risk and Insurance - Issues and Practice, 46.

[98] Woods, D. W., Bohme, R. and Moore, T. (2021). The county fair cyber loss distribution: Drawing inferences from insurance prices. Digital Threats: Research and Practice, 2.

[99] Wu, M. Z., et al. (2021). Modeling multivariate cyber risks: Deep learning dating extreme value theory. Journal of Applied Statistics, 50.

[100] Xu, M. and Hua, L. (2019). Cybersecurity insurance: Modeling and pricing. North American Actuarial Journal, 23.

[101] Yang, Z., et al. (2020). Premium calculation for insurance businesses based on cyber risks in IP-based power substations. IEEE Access, 8.

[102] Zeller, G. and Scherer, M. (2021). A comprehensive model for cyber risk based on marked point processes and its application to insurance. European Actuarial Journal, 12.

[103] Zeller, G. and Scherer, M. (2024). Is accumulation risk in cyber methodically underestimated? European Actuarial Journal, 14.

[104] Zhang, Q., et al. (2018). A fuzzy probability Bayesian network approach for dynamic cybersecurity risk assessment in industrial control systems. IEEE Transactions on Industrial Informatics, 14.