Quantified Security is a Weak Hypothesis

A critical survey of results and assumptions

Vilhelm Verendel
Department of Computer Science and Engineering
Chalmers University
vive@chalmers.se

ABSTRACT

This paper critically surveys previous work on quantitative representation and analysis of security. Such *quantified security* has been presented as a general approach to precisely assess and control security. We classify a significant part of the work between 1981 and 2008 with respect to security perspective, target of quantification, underlying assumptions and type of validation. The result shows how the validity of most methods is still strikingly unclear. Despite applying a number of techniques from fields such as computer science, economics and reliability theory to the problem it is unclear what valid results exist with respect to operational security. Quantified security is thus a weak hypothesis because a lack of validation and comparison between such methods against empirical data. Furthermore, many assumptions in formal treatments are not empirically well-supported in operational security and have been adopted from other fields. A number of risks are present with depending on quantitative methods with limited or no validation.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection; H.1 [**Information Systems**]: Models and Principles; C.4 [**Performance of Systems**]: Modeling techniques

General Terms

Measurement, Reliability, Security, Verification

Keywords

Quantitative security models, Security metrics, Validation

1. INTRODUCTION

Much of the work on *quantitified security*, the quantitative representation and analysis of computer and information security[1], is motivated by variants of the following idea: we can not control what we can not measure. To know how well security requirements are met, a major challenge is to provide precise knowledge of security properties in relevant operational settings. One way to approach this problem is to attempt quantification. Decision-makers need information about the utility, such as the reliability or estimated costs and impacts, of different security options to make appropriate decisions. Quantification has been suggested as a solution to such needs. Mostly during the last decade it has been claimed in scolarly literature and by leading standards organizations that such quantification is not only possible but also, in a wide sense, beneficial [SH00], [BMG01], [But02], [Sch02], [SBS+03], [AHP+04], [AL05], [Pay06], [CSS+07], [ANS08] and sometimes even necessary for good security management. This suggests that good quantification of security comes with requirements, most important that the quantitative methods need to be valid.

Thus, a large amount of work claims that we need a quantitative view of security. However, such claims depend on the following hypothesis: *security can correctly be represented with quantitative information*.

While there is no way to definitely accept or reject such a general statement, this paper now proceeds to evaluate the current strength of this *quantification hypothesis* in detail. The keyword in the hypothesis is *correctly*, which will depend on existing efforts of hypothesis evaluation. Do quantitative methods work? Do they correctly describe security? Answering this requires studying a number quantitative methods. To that end we survey a large part of the work we found by on quantified security between 1981 and 2008. The different methods found are in a surprising variety of forms: from heuristic "rules of thumb" for counting observations, to the use of professional experts for quantitative judgement and assessment, to proposing econometric quantitative indicators, using quantified security information to test hypotheses with statistics, to formal analysis to model security in system reliability settings. These different approaches to security quantification are examined, and 90 papers are classified with respect to security perspectives, targets of quantification, underlying assumptions and type of validation effort by e.g. empirical methods. The goal is to evaluate the work by critically assessing validity of the different methods, and whether they support the quantification hypothesis or not. The work is mostly academically published[2].

The result from the survey shows how there exists significant work for quantified security, but there is little solid evidence that the methods represent security in operational settings. A number of techniques from e.g. computer science, economics and reliability have been applied to quantify security, but after around 20 years of work on the subject little is known about the validity of the results.

[1] In the rest of the paper, referred to as *security*

[2] Likely to have undergone peer review. A smaller number of cases are e.g. unpublished, but publically available technical reports or standards.

NSPW'09, September 8–11, 2009, Oxford, United Kingdom

Without proper validation of such quantitative methods there are a number of risks when it comes to operational usage: from economic, engineering and usability standpoints. We discuss problems that makes quantification hard and finally consider what may be required for progress in the field.

2. BACKGROUND

Here, terms and concepts are defined for the rest of the paper (despite that ambiguity exists in parts in the literature under review, e.g. with security measurements and metrics).

2.1 Weak hypotheses

To make the concept of a *weak hypothesis* (our terminology) more precise, we relate to Karl Popper's model [Pop59] of scientific knowledge particularly in the empirical sciences. In this view descriptive (e.g. quantitative) methods that attempt to represent empirical facts are seen as *hypotheses* that can be either incorrect or correct to a varying degree. While many methods of generating hypotheses are important, the ultimate and crucial way to learn about the correctness of a hypothesis is by challenging empirical tests. If hypotheses are successful in describing outcomes of experiments, e.g. by repeated large-sample tests, they get *corroborated*. Alternatively, hypotheses can also be *falsified* by inconsistent evidence and anomalies. The latter requires replacing or modification of hypotheses.

In the light of Popper's view above, we will call a hypothesis *weak* if it lacks clear tests of its descriptive correctness. This denotes a hypothesis (here a quantitative method) where too little is known about its correctness to call it corroborated or falsified. This ultimately depends on a lack of empirical tests or unclear evidence.

2.2 Measurements, metrics and models

Security measurements, metrics and models are here related to each other in the following way. A *measurement* is made by observing the outcome of an event using some appropriate method to collect the result into data. A *metric* assigns such data onto some kind of scale in order to correctly represent some security attribute of a system under consideration [BF08]. This scale is in our case quantitative[3] in order to provide precise evaluation of systems. For valid assignments, this allows correct and precise assessment and comparison of systems using e.g. the absolute values or number ordering. Furthermore, the idea of a security *model* in our context is to provide a formal representation (e.g. sets of equations) that corresponds well to security for systems under consideration. A valid model can then be used to derive quantities of interest using appropriate parameters and data.

Models are required for quantification as soon as there is a non-trivial relationship between possible measurements and the attribute that one wants to quantify: data from potentially imprecise or uncertain measurements needs to be related to some definition of security. In the light of this, building and validating quantitative models for security is of crucial interest when one requires or claims accuracy in describing security.

It is clear that metrics and models for security need validation: using correct assumptions or success with describing actual evidence. Otherwise they may have limited utility: in the worst case a model of the modeler(s) understanding rather actual systems under consideration. For measurements, it is hard to make clear distinctions in some cases: measuring is dependent on assumed models of the events they are meant to observe. An example is security intrusion detection: events of interest to observe are intrusions (with known limitations and uncertainty[Axe00]), but e.g. counting such events in practice blur the line between measurements, metrics and models. Validation is emphasized later in the paper, but mostly ignoring these subtleties since they are unnecessary for the main point.

[3]This paper considers work with quantification mostly using the ratio scale, which allows comparisons and certain other properties.

3. OPERATIONAL SECURITY

We are mainly considering quantification of *operational security*: security for systems in realistic environments (e.g. the Internet, the infrastructure of an organization, or any realistic interaction with non-trivial threats). Computer and information systems are constantly being exposed to a variety of threats, potentially leading to attacks and violation of security policies. This section presents a simple conceptual model of operational security and some characteristics that make modeling of security a challenging problem.

3.1 Basic components

This paper uses a minimal and coarse conceptual model of security in order to easily incorporate different lines of research into the same framework. The survey views operational security as having at least the following components

- **Systems**: technical/structural components, security controls and users.
- **Threats**: active and passive agents (ranging from people to code and random failures) capable of attack and violating security properties of a system.
- **Vulnerabilities**: system properties that allow realization of attacks.

The structure and interaction between these (often interdependent) components leads to a number of events over time: a threat needs to interact with a vulnerability in a system to potentially turn an *attack* (attempt to use vulnerability) into a successful security *breach* (with respect to a systems security policy). Breaches have further *loss impact* (that can often be understood in economic terms) as well as further secondary events such as structural change of systems or threats. Systems often contain a number of components that are security controls that are the subject of quantification. As noted above, this work deliberately ignores finer and detailed descriptions of systems except for short description in Tables 1, 2 and 8.

3.2 Characteristics of Operational Security

A number of properties make operational security a challenge to understand and model, compared to some other problems approached with quantitative methods. It is characterized by the following properties, among others:

- **Dynamics**, with systems and threats that are adaptive and learning from interaction. Threats with learning may adapt to security mechanisms in systems, and systems themselves may adapt to threats and security events. When vulnerabilities have been used, threats may target novel types - or similar, but unknown, vulnerabilities. Security decisions often depend on interplay and deliberate tradeoffs [KH03, LW05, BOS07, GKMR07] for the different agents that are involved. Thus, even in a fixed environment of systems and threats the behavior and patterns of security events may change over time.
- **Low stationarity**, with systems, threat environments and vulnerabilities that may change rapidly [Sch04]. This may have

impact on the robustness of security properties. An example is software systems: composed of parts that are easy to update, patch and re-engineer. This has its benefits in terms of flexibility, but makes systems more complicated. Furthermore, novel threats and attack methods may appear quickly and change the threat environment. An example is the Internet: spatial and temporal distance often set no barriers for threats and do not prevent quick changes. Threats may not only learn and improve over time, but also disappear or arrive which is hard to observe [Ozm05]. Such potential changes of many basic components may lead to changing patterns of security events (compared to the dynamics in a given system).

- **Economics**, with self-interested agents (attackers as threats, but also defenders and other decision-makers with interests) seeking to achieve their goals where there is conflict of interest [And01, Bie04, LW05, AM07]. Security attacks are often realized by planned actions rather than random faults [MKF03, GKMR07]. This means that failures may occur and be realized differently than in physical systems that are the target of constant physical stress.

- **Dependence**, with intentional and strategic threats (usually people) that do not act seemingly independent and randomly (such as churn in physical systems, e.g. hardware reliability). Security failures are often highly *directed, designed faults*. Attackers have the capability to plan and launch sequences of attacks that are highly correlated (in e.g. time or targets) against different systems and components. Evidence shows that many processes in security contain events that are not independent [MKF03, NST04, Bie04, Ozm07, KDA$^+$07]

- **Uncertainty**, with most decisions in security have to be made with uncertainty and limited information about consequences of actions [Ver08]. Many factors cannot be directly observed when making decisions: measurement of remote threats can be hard and uncertain, and controlled measurements such as intrusion detection has limits [Axe00] and is far from perfect. This requires agents in operational security to decide with uncertainties.

The properties above will be used for framing the analysis in the following sections.

4. SURVEY METHODOLOGY

The survey presents (to the authors best knowledge) a significant and representative overview of work between 1981 and 2008 containing

1. **Security models**: formal representation of operational security aimed to represent quantitative information of interest

2. **Security metrics**: quantitative measurement-based indicators for various targets in operational security

3. **Security frameworks**: general methods for quantitative analysis of security (not necessarily specific models or metrics), proposing frameworks and mechanisms for quantification

The rest of the paper will sometimes simply refer to these as *quantitative methods*. Thus, our criteria are broad, but this is a natural consequence arising from the number of different approaches that have been taken in attempts to quantify security. Ultimately, we claim that the selected material is related to the quantification hypothesis that maybe allows an even broader perspective.

In practice, the following method was used to select, classify and analyze the material:

1. The initial material consisted of 140 documents, either peer-reviewed research articles, technical reports or part of the standards literature that is emerging in the area. These were found by exhaustive search: articles were picked from from various sources and databases in attempts to match the three criteria of quantified methods above. The limiting factor was the effort the author could spend on collection.

2. The work found have appeared since 1981, but there seems to have been almost no mentioning of the problem before around 1990. The material that was considered is limited to between 1981 and 2008[4].

3. Work clearly diverging from operational security was not taken into account: such specific work which was discarded is quantitative analysis of cryptographic algorithms, information flow, or e.g. privacy metrics that offer highly specific analysis rather than of computer and information security in operational settings (as in Section 3).

4. A number of articles were discarded because of redundancy: the same or very similar ideas appeared in more than one publication, often by the same (or common subsets of) authors that have gradually extended an idea.

5. The taxonomy of perspectives, targets and assumptions were chosen by manual study of the original works.

Thus, the selection depends on judgement of a number of cases that may be unclear and not contained in the analysis, as well as limits to effort that could be spent. However, it appears to the author that most developed work with clarity and complete presentation has been selected, as well as all the major approaches that in the literature. The 90 articles selected are found by timeline in Tables 1, 2.

4.1 Taxonomy

The material was analyzed with respect to different variables, in order to understand which methods and support exists for the quantification hypothesis. With big diversity, there are a number of problems to understand and differentiate the quantitative methods that have been proposed. First, validity of a method depends on a clear definition or perspective of security. What is to be quantified is thus a matter of definition, which may influence how validation should be performed[5]. Second, it is important to know what security components, attributes or events that are considered as *target* for quantification. Third, to understand when a method is valid it is necessary to examine what assumptions are made for specific systems under consideration. It is important not to overgeneralize under specific assumptions. Finally, which different methods have been used to evaluate and validate results have to be understood.
In order to clarify these points, the classification is with respect to the following variables

- **Perspective**: from which conceptual viewpoint the approach to security is taken. This depends on explicit or implicit motivation and in which scope the general idea of security is presented.

[4]Systematic study of the area seems to have started in the early 1990s and started to get increased academic momentum during the last decade.

[5]A view of security in qualitative terms of e.g. Confidentiality, Integrity and Availability may suggest different validation methods than e.g. security as a probabilistic guarantee of system attributes.

- **Target**: targets with, or related to, operational security that are considered for desciption with quantitative methods.
- **Assumptions**: specific assumptions that were used, explicit or implicit ones. Important working assumptions that were used to model targets or motivate validation and quantitative methods.
- **Validation**: which kind of work that was used for validation, generally attempting to support or evaluate the quantitative method under consideration. Which kind of methods that were used to support claims and results.

4.2 Perspective

Perspectives were classified as

- **CIA (CIA)**: the classic viewpoint of Confidentiality, Integrity and Availability.
- **Economic (ECO)**: economics and risk analysis. This typically consists of analysis in terms of different self-interested agents or risk and trade-offs with various consequences.
- **Reliability (REL)**: reliability and dependability theory. This typically means to consider stochastic processes and probabilistic analysis for the rate of failure events in a system.
- **Other (OTH)**: various other techniques, often from computer science. One example is the extensive use of different graph models. Further specified among the keywords.

4.3 Target

The targets for quantified methods were classified as

- **Economic (ECO)**: economic efficiciency, incentives, impact and risk of security events and threats.
- **Framework (FRA)**: framework regarding how to develop and select quantification method (models or metrics)
- **System (SYS)**: components and their structure in the system under consideration, and how they relate to security.
- **Threat (THR)**: the threat (active or passive), attacker motivations (rewards) and actions (attacks, breaches).
- **Vulnerability (VUL)**: existence or appearance of system vulnerabilities.

4.4 Assumptions

The following assumptions were found and classified as

- **Independence (IND)**: that random events in systems occur with *probabilistic independence* of each other. In formal security models, this has mostly been used to assume that complicated systems and sequences of events can be split into simpler independent parts. By assuming independence, it becomes easier to derive quantitative properties in the equilibrium behavior of systems in stochastic models. Two techniques assuming independence are common: Markov modeling (system states, in discrete or continous time) and arrival processes (how sequences of events occur). First, Markov models use the "memoryless" assumption: given a current state of a system, the following realizations are in probability conditionally independent of the past. Independence is often implicit in the use of random variables being exponentially distributed in time. Second, modeling arrivals (e.g. vulnerability apperance) typically with the Poisson process: again, with security events occuring independently of each other (uncorrelated and rarely in bursts).
- **Rationality (RAT)**: that the different agents (mostly human) will act rationally in uncertain situations, or in the case of experts - having rational judgement in providing information. Assuming rationality for formal models allows considering problems such as attacker modelling as formal optimization problems. For methods that depend on expert input, assuming that provided information is always the accurate description of some target.
- **Stationarity (STA)**: that quantified system, threat or vulnerability properties are *invariant* over time or between environments. Here, when quantitative models are proposed with numerically fixed parameters without motivation. Generally, this also holds if parameters for a general method are presented as fixed from referring to limited (or no) empirical instances. This assumes that systems (e.g. software), threats (e.g. attackers) or vulnerabilities do not significantly change characteristics over time[6].
- **Additive (ADD)**: that quantitative information from system components can be composed in *an additive* manner to correctly describe a system from a quantitative description of its parts. This is sometimes used when modelling a hierarchical system or attacks in terms of their smaller and simpler parts.

4.5 Validation

Validation methods and work were classified as having

- **Hypothetical (HYP)**: hypothetical examples, having unclear relation to actual phenomena and degree of realism.
- **Empirical (EMP)**: empirical methods, such as systematic experimental gathering of data from operational settings.
- **Simulation (SIM)**: simulation methods of some target.
- **Theoretical (THE)**: formal or precise theoretical arguments to support results. This includes e.g. relating a metric or model to what attribute properties it is meant to represent and measure [7].

The validation variable does not directly attempt to reflect correctness or quality of results. It attempts to describe the kind of work and the methods that have been used in work under consideration.

4.6 Keywords

Additionally, the material was described with a few relevant chosen keywords for target, scope and validation of the work.

4.7 Previous surveys

A few previous surveys have been found relating to security modelling and quantification, but specific to the different perspectives under consideration. The two most relevant are [NST04] which considers models mostly from system reliability, and [VMP04] is specifically considering explicitly named security metrics. While being a few years old and more specific, there are no obvious inconsistencies with respect to our findings.

[6]It is clear that the assumption of stationarity can be made more general to relate to models and methods: here, it is considered in a more limited sense.

[7]This is mostly conceivable in the following sense: recall that a metric needs validation in its relationship to a given attribute. The challenge is to relate a metric to a property of the target attribute.

	Perspective	Target	Assumptions	Validation	Keywords
[CS81]	ECO	ECO,FRA,SYS	STA	HYP	Cost effectiveness of security controls, risk management, organizational
[Str90]	ECO	ECO		EMP	Organization,Security efficiency, person survey
[LBF+93]	REL	FRA			Applying reliability to security, attacker effort
[AFB95]	OTH	SYS	ADD		System Vulnerability Index (SVI)
[DDK96]	REL	SYS,THR	IND,STA	EMP	Privilege graph, Markov chain/Petri nets
[VGM+96]	REL,OTH	SYS,THR		SIM	Adaptive vulnerability analysis, fault injection, time to intrusion
[JO97]	REL	SYS,THR	IND	EMP	Attack effort, breach process, university students
[WW97]	CIA	FRA	ADD		Metrics for CIA, composing metrics
[SPG98]	ECO,OTH	SYS,THR,VUL	RAT	HYP	Attack graph, network vulnerabilities, risk analysis, depends on database, attacker profile, 6 scenarios
[ODK99]	REL	SYS,THR	IND,STA	EMP	Fixed vulnerabilities, one large system over two years
[Sch99]	ECO,OTH	THR,SYS		HYP	Attack trees, propagating values, needs expert input, several examples
[Mer99]	ECO	ECO,FRA,SYS	STA	HYP	Broad method for quantitative risk analysis, depends on experts, risk scenarios
[SW00]	CIA	THR			Adversary work factor, data not published
[SH00]	ECO	ECO,FRA	RAT	THE,HYP	Decision modeling, optimal safeguards, monetary metric, organization
[Low01]	ECO,OTH	THR	ADD	SIM,HYP,EMP	Adversaary planning, attack trees, model parametrized by experts, composite metric, red team 5 days
[BAMF01]	ECO,OTH	THR		EMP	Exploit trends, model incident rate, trends, CERT data, self-reporting
[SdBF+02]	CIA,ECO,REL	FRA	ADD		CORAS, model-based risk assessment, component-based software, UML
[HMOS02]	ECO	FRA,THR	RAT	HYP	Analysis of security using game theory
[SGF02]	CIA,ECO	FRA,ECO		HYP	NIST, Risk scale, neccesary actions, cost-benefit analysis
[SHJ+02]	OTH,ECO	FRA,SYS,THR	IND	EMP,HYP	Auto-attack graphs, model-checking, evaluate controls, 4 exploits
[MGPVT02]	CIA,REL	SYS,THR	IND	SIM,THE	Semi-Markov, unknown parametric fit/simulation correctness
[SH02]	ECO	ECO,FRA		HYP	How much spending is enough, risk management, ALE, decision analysis
[Sch02]	ECO	FRA,ECO	RAT	THE	Cost to break-metric, using vulnerability market
[But02]	ECO	ECO,FRA	RAT	HYP,EMP	Cost-benefit, priorities, needs experts, survey and case study
[SBS+03]	ECO,OTH	FRA,ECO	RAT		NIST metrics, regulatory, goal-related, improve security investment decisions
[SH03]	ECO	ECO,FRA	STA	HYP	Risk analysis, weighing asset relevance, ALE, business model
[ADSW03]	ECO	ECO,FRA		EMP	OCTAVE, needs experts, organizational, small organizations but unclear results
[HPW03]	ECO,OTH	SYS,THR	STA,ADD	EMP	Attackability, state model, test with security bulletin, unclear result, fixed weights
[AB03]	ECO	ECO,SYS,FRA	RAT	THE	Network intrusions, trade-offs, game theory, admin decision and analysis
[MKF03]	CIA,REL	FRA,SYS,THR	IND	THE,HYP	Merging random/security faults, designed faults, stochastic algebra
[MT04]	CIA,REL	SYS,THR	IND	THE,SIM	Attack-response graph, time to security failures, Markov model, guessed parameters
[MGPVT04]	CIA,REL	SYS,THR	IND		SITAR, mean time/effort to security failure, software impact, unknown fit of parametric model
[SLN04]	REL	SYS,THR		THE,SIM	Model-based testing, attack hardness, metric estimation, importance sampling, heuristical simulation
[DLK04]	ECO	THR,SYS	RAT	HYP	Behavior-based attack graphs, risk analysis, Bayesian, unclear attack optimization
[Sch04]	ECO	FRA,ECO,THR		HYP	Econometric models for security risk, safeguard efficiency, security strength
[MW04]	OTH	ECO,SYS,THR		EMP	System attack surface, attack classes, applying relative metric on Linux versions, unclear goal
[Ozm05]	REL	VUL		EMP	Reliability growth models, limited accuracy, OpenBSD vulnerability data
[MBFB05]	ECO	THR	STA,ADD,IND	EMP	Time to compromise, risk efficiency, SCADA system, known vulnerabilities
[DPP05]	CIA,ECO	SYS	ADD	HYP	Risk, risky trust, software components, design
[Nic05]	ECO,OTH	FRA		HYP,SIM	Modeling and Simulation in Security Evaluation, impact assessment, simulation methods
[KS05a]	ECO,OTH	ECO,SYS	IND,RAT,STA		ISRAM, risk analysis and software, organization, needs experts, case study
[MW05]	OTH	SYS		HYP	Attack surface metric, attackability, attacker effort, source code, maybe incomparable metrics
[GMT05]	REL	SYS,THR	IND	THE,HYP	State space, software, security-failed states, Markov, safety, controllability, optimal policy
[SKH05]	ECO,REL	THR	RAT	THE,HYP	Expected attacker behavior, stochastic game theory, cost/reward tradeoffs
[CH05]	ECO,OTH	SYS,ECO		EMP,HYP	Fuzzy set theory, network security, unknown military expert parameters/ranking
[Mcd05]	REL,OTH	SYS,THR,VUL	IND	HYP	High-consequence systems, competent attacker potential, survivability, stochastic process algebra
[KS05b]	OTH	VUL,SYS		SIM,HYP	Vulerabilities, security level estimation, design/exploitation stages, metric hierarchy
[LW05]	ECO	THR,SYS	RAT,STA	HYP,THE	Game strategies, network security, stochastic games
[LZY05]	ECO	ECO,FRA	RAT	THE,SIM	Games for inference of attacker intent/objective/strategy, case study, DDoS attacks
[PJAS06]	OTH	SYS,THR,VUL		HYP	Weakest-adversary metric, known vulnerabilities, attack graph, requirements algorithm, penetrability
[Pay06]	ECO	FRA			SANS, 7-step framework, decisions, useful metrics for improvement and value
[Waw06]	ECO	FRA,ECO		THE	Security threat risks versus cost of security measures
[PPN06]	ECO	FRA	STA,ADD		Risk management, 25 metrics, security goal performance metric, organization
[HHH06]	ECO,OTH	FRA,ECO,SYS	ADD,STA	HYP	XMASS/crossroads framework for complex networks, fixed weights, applying methods, unclear goal
[KS06]	OTH,ECO	SYS,FRA,THR		EMP,SIM	Attack graph complexity, network security metrics, simulation, experts, graph from real network
[MBFB06]	OTH	SYS,THR	IND	SIM	Risk reduction measure, known vulnerabilities, compromise graph, time-to-compromise, SCADA
[Bie06]	ECO,REL	FRA,THR	RAT		Combining reliability and game theory
[LS06]	ECO	THR	IND	THE,EMP	Port-scans, potential loss measure, stochastic model, security drift, university port-scans
[YCL06]	ECO	ECO			Graph Model, virus, risk assessment, network security, genetic algorithm, investment optimization
[WT06]	OTH	SYS	ADD	HYP	Composed systems, aggregating component measures, algebraic and/or/mean,
[SHK06]	ECO,REL	SYS,THR,FRA	IND,RAT	HYP	Estimated attacker behavior, operational measures, zero-sum game theory, Markov model
[BFP06]	ECO,OTH	ECO	RAT	HYP	Defense trees, evaluating security investments, attacker return on attack, control decisions
[ZWW06]	OTH,ECO	SYS		SIM	Network survivability, decision matrix, relational analysis, entropy difference, unclear success
[NR06]	OTH	SYS		EMP,HYP	Metric tree, unspecified metrics, decisions, dependency graphs, applied to VoIP, unclear result
[BLP+06]	ECO	ECO,THR	RAT	THE,HYP	Multi-parameter attack trees, choosing measures, game theory, hypothetical company

Table 1: Classification of material (1981-2006). Perspective: CIA = Confidentiality/Integrity/Availability, ECO = Economic, REL = Reliability, OTH = Other. Target: ECO = Economic, FRA = Framework, SYS = System, THR = Threat, VUL = Vulnerability. Assumptions: ADD = Additive, IND = Independence, RAT = Rationality, STA = Stationarity. Validation: EMP = Empirical, HYP = Hypothetical, SIM = Simulation, THE = Theoretical.

	Perspective	Target	Assumptions	Validation	Keywords
[MSR07]	CIA,ECO	FRA,VUL	ADD,STA	EMP,HYP	CVSS, vulnerability score, prioritize risk, fixed weights, applied to 3 vulnerabilities
[CSS+07]	ECO,OTH	ECO,FRA			NIST, measurement guide, security efficiency, organizational decision-making, 19 measures
[Hau07]	ECO,REL	FRA,ECO,SYS,THR	RAT	THE,HYP	Infrastructures, reliability, game theory, dependence, investment optimization, optimization, examples
[WSJ07]	OTH	FRA,SYS,ECO		HYP	Network security measurement, attack graph, imprecise measures risk
[RGH07]	ECO	FRA,SYS,THR,VUL			Comparing risk assessment methods, SCADA, compromise graph, vulnerability trees, risk reduction
[Ozm07]	REL	VUL,FRA		EMP	Improving vulnerability detection models, assumptions, several datasets, independence not working
[KMR07]	REL	VUL,SYS		EMP	Vulnerability discovery rate, risk assessment, different software versions, shared code, apache/mysql
[MKW07]	ECO,OTH	FRA,SYS,THR,VUL	ADD	EMP	Attack surface metric, formal model, I/O automata, similar software systems, anecdotal evidence
[MTMW07]	OTH	FRA,SYS		EMP	Attack surface metric, similar systems, metric application, expert perception, security bulletins
[MYY+07]	ECO,OTH	FRA,ECO,VUL	ADD,STA		Network security, hierarchical analysis, attack graph, risk of different levels, policy development
[KDA+07]	OTH	THR,SYS		EMP	Modeling attack processes, regression, mixture model, showing dependence, 35 Internet honeypots
[SHK07]	CIA,ECO,REL	ECO,SYS,THR	RAT,IND	THE,HYP	Integrated security/dependability assessment, security measures, real-time, Markov, game theory
[SCHB07]	ECO	ECO,FRA,SYS,THR	IND,RAT	HYP,SIM	Comparing attacks, network situational awareness, quantitative prediction, simulated network attack
[BDP07]	ECO	ECO,VUL,THR,SYS	RAT	HYP	Strategic games, defense trees, risk analysis, return on investment, measure countermeasure efficiency
[JW07]	ECO,OTH	ECO,THR	RAT,ADD	THE,HYP	Multi-parameter attack trees, interval estimates, economic security level, hierarchical assessment
[AAP07]	ECO	FRA		HYP	Automated Model-Based Risk Analysis, patterns, vulnerabilities and metrics, decision-making
[BM07]	ECO,OTH	FRA	ADD	EMP	14 metrics, cyber control systems, driving decisions, 7 security dimensions, case study
[HSHJ08]	ECO	FRA		SIM	Security patterns, composing metrics, control metrics, software, ATM case study
[ANS08]	ECO	FRA,ECO			ANSI, guide in quantification of cyber risk, 50 questions to evaluate, experts, policy roles, ALE
[LJ08]	REL,ECO	SYS,THR	IND	HYP	Estimating time-to-compromise metric, cost ratio, company case study
[WIL+08]	OTH	VUL,SYS		HYP,THE	Attack-graph probabilistic metric, dependency cycles, metric algorithm, hypothetical examples
[AM08]	REL	VUL		EMP	Vulnerability discovery models, 6 models, operating systems, some models relatively better
[YSH+08]	OTH	SYS	ADD	HYP	Software architecture, pattern metrics, overall indicators, aggregation algorithm, hypothetic case
[Hul08]	ECO	ECO,THR,VUL	IND	THE,SIM	Value at Security Risk, investment decisions, communicating risk, transforming to economic metrics
[GMMS08]	ECO,REL	ECO,VUL		EMP	Assessing network vulnerability, comparative measures, case study:Internet backbone

Table 2: Classification of material (2007-2008). Perspective: CIA = Confidentiality/Integrity/Availability, ECO = Economic, REL = Reliability, OTH = Other. Target: ECO = Economic, FRA = Framework, SYS = System, THR = Threat, VUL = Vulnerability. Assumptions: ADD = Additive, IND = Independence, RAT = Rationality, STA = Stationarity. Validation: EMP = Empirical, HYP = Hypothetical, SIM = Simulation, THE = Theoretical.

5. ANALYSIS

To evaluate and seek an answer to our main question - what the current strength of the hypothesis about quantified security is, we need to assess the empirical support as outlined in Section 2.1. It is important to acknowledge that empirical work can support quantitative methods *directly* by evaluating specific methods, but also *indirectly* by supporting previous results and assumptions for validation using simulation and theory. We find that most assumptions and methods are *weak* and not well-supported.

To examine this, three questions are considered. First, where effort has been made - where different perspectives meet and where they diverge. *What have different perspectives focused on?* Second, what assumptions that are used in different methods and models. Are they reasonable, given available evidence as well as operational security as described in Section 3.2? The available support of formal assumptions affects how well models without empirical evaluation likely describe operational security. *Are common assumptions well-supported and reasonable?* Finally, considering which empirical work exists that supports and validates methods to quantify security (which, reminding ourselves, is a matter of definition). *What empirical work was done and results were achieved?*

5.1 What different perspectives focused on

The survey data allows a rough summary of where different perspectives have been used - which may indicate where there exists more limited quantification effort (or where novel approaches are needed). At least three issues are relevant here. First, what security targets have been in focus for work taking different perspectives (Table 3). Second, what kinds of validation that has been made by different perspectives (Table 4). Finally, which kind of validation methods has been attempted in order to develop quantitative methods for the different targets (Table 5).

In Table 3 we see that the CIA perspective on quantified security is underrepresented, and that a majority of work at least contains or has grains of an economic perspectives (this depends on many papers typically providing economic and decision-making goals for developing quantitative methods). This is interesting, since CIA is the typical frame in which applied security is viewed. The difference likely depends on that other perspectives (such as risk and reliability models) have previously been developed and applied to different problems of established quantitative nature, which has forced the focus of tangible properties and interesting parameters. However, for the CIA perspective which belongs to traditional security a standard quantitative representation seems to have been lacking, reflecting that CIA is qualitative in nature. This may suggest that the field of security needs a shift of viewpoint before quantification becomes natural.

A similar pattern of underrepresentation may be present for the reliability perspective, with the exception for threat and vulnerability analysis (which later examination shows has made a concentrated

Perspective	ECO	FRA	SYS	THR	VUL	Total
CIA	2	5	6	6	1	11
Economic	33	35	23	26	8	59
Reliability	3	7	16	17	6	24
Other	10	11	25	15	7	34

Table 3: How different perspectives have focused effort on various security targets. Targets are economics (ECO), framework (FRA), system (SYS), threat (THR) and vulnerability (VUL).

Perspective	EMP	HYP	SIM	THE	None	Total
CIA	1	5	2	4	4	11
Economic	16	31	8	13	11	61
Reliability	8	7	4	8	4	24
Other	12	18	7	2	5	34

Table 4: How different perspectives have used different types of validation effort. Validation methods are empirical (EMP), hypothetical (HYP), by simulation (SIM), theoretical (THE).

effort much bigger than its proportion at tackling the problem). The frequency of the other (OTH) perspectives shows there is room for improvement in this taxonomy.
Furthermore, Table 4 summarizes the validation methods that were found. This indicates that the reliability perspective appears to be the most well-defined area with most empirical work[8] (and relatively lowest degree of hypothetical examples). The CIA perspective is clearly underrepresented in attempting to validate using empirical work. Again, the other perspective shows that there is room for improvement in the classification. The same pattern seems to appear when considering Tables 3 and 5 - a focus on systems, threats and vulnerabilities seems to be relatively easier targets for empirical work as well as hypothetical examples.

Target	EMP	HYP	SIM	THE	None	Total
Economic	6	18	3	10	6	33
Framework	9	19	5	7	11	41
System	13	27	9	10	4	48
Threat	13	21	9	13	4	42
Vulnerability	7	7	2	2	2	16

Table 5: Attempts of validation for different targets. Validation methods are empirical (EMP), hypothetical (HYP), by simulation (SIM), theoretical (THE).

Target	ADD	IND	RAT	STA
Economic	3	4	14	6
Framework	8	4	11	7
System	7	16	10	9
Threat	5	18	13	6
Vulnerability	3	2	2	2
Total	16	19	21	14

Table 6: Assumptions used for different targets. Assumptions are additive (ADD), independence (IND), rational (RAT) and stationarity (STA). Independence and rationality are most representative especially for systems and threats.

5.2 Are common assumptions reasonable?

This section examines common assumptions that were found in some generality. It is often the idea of a model to simplify and work with assumptions, but without other validation the correctness of a descriptive model depends on the support of those assumptions. An overview of the assumptions (described in Section 4.4) that were found in the material are summarized in Table 6 and shown in Tables 1, 2. The method to analyze the assumptions is comparing them to the description of operational security in Section 3.2, as well as other empirical findings.
The result shows that the assumptions we have classified either have conceptual differences with respect to operational security, or that there is evidence speaking against them:
Independence: with assuming probabilistic independence. Depending on how one views security systems, threats and vulnerabilities there are theoretical and empirical problems. There is empirical evidence based on statistical tests that a number of security events are not always modelled well by assuming probabilistic independence in processes with exponentially distributed interarrival times. These include vulnerability appearance [Ozm07] as well as e.g. attacks over the Internet [KDA+07]. Conceptually, using independence to model sequences of actions seems to contradict the assumption of threats capable of planning and coordianate highly correlated (or dependent) attacks. The same argument can be made for separate system components.
Rationality: with assuming rational agents that act optimally. The problem here lies in mostly the threat and system users: can they be assumed to be rational in conflict scenarios and in judgement? In technical security, the degree of agent rationality has been little evaluated and remains unclear. However, the available evidence from other fields (involving risk and uncertainty) suggests that this is far from obvious for decision-making in operational security [Sch07, Ver08].
Rationality is closely related to what attacks and what agents that can constitute threats and other actors in operational security. When it comes to intelligent adversaries, it is mostly assumed that the threat are humans rather than e.g. automated software agents (who can not yet, in many cases, be expected to have the same level of foresight and planning). A large body of empirical work shows that while humans can be sophisticated in planning and foresight they are far from always perfect from making optimal decisions. The idea of rationality has been under empirical attack for three decades, but it is unclear what will come out of it for formal models. Examples of such work is found in the field of *bounded rationality* (starting from [Sim55]) which studies systematic limits in human decision-making. Likely, this may have some relation to how agents in security trade-off risks and costs using decision *heuristics* [KST82, Kah00] rather than performing computational optimization. The question of rationality is thus whether it is a realistic approximation to agent behavior in security, but too little is known about how decisions are made.
Stationarity: with assuming that some (measured or quantified) properties in modeled targets are constant between model targets or over time. This mostly relates to how risks can vary or not between different systems and threats. It has been as pointed out, perhaps most explicitly by Schechter [Sch04], that many trends in computer and information systems undergo quick changes that makes systems little stationary. Compared to the physical world with physical instruments, where trends of threats usually change dynamics and preference slowly over time [Sch04], systems may change quickly when it comes to the use, structure and threats of typical computer and information systems. For further examples, see Section 3.2.
This survey found a number of places where stationarity is assumed: a number of methods are presented with already fixed numerical values in their models, without further motivation. This either assumes that the parameters will not vary between systems, or in the case of predictive methods over time. Among other places, this was found in several risk-based approaches: using fixed parameters (usually without empirical motivation) to weigh together risks and quantified information.
This general volatility of security properties may form an environment that is hostile for forward-looking, typically statistical, methods. While statistical methods may be used to separate previous noise well from previous data, this leaves the question open for prediction that is inductive in nature. This, of course, does not rule out that assumed regularities actually exist - but presentation for evidence of that is typically lacking. Few studies have been made regarding to the stationarity of security environments, but one can expect some environments and systems to quickly change [Sch04]

[8]While the Other perspective has relatively more empirical work, this denotes a number of different approaches that this survey has not characterized: a more fine-grained classification would show this.

which may make tests and evaluation hard [Ozm05] and limit the utility of quantitative methods for forward-looking decisions. This seems to require quantitative methods to be adaptive to information that appears on-line in a changing environment.
Additive: with methods using addition to produce a quantitative description of systems from their smaller parts. This is often without motivation when aggregating numbers to represent quantification of compound system components. While there seem to be few studies of this specific question, it still deserves questioning since it completely rules out mutual dependence (e.g. perhaps multiplicative effects?) between components in some of the models where it occured.

After having considered these assumptions it is clear that using them to model targets in operational security is far from obious. First, because of their lack of conceptually describing operational security (presented in Section 3.2). Second, because there exists contradictory evidence; at least to some degree as shown above.
However, it could be argued that these assumptions are just necessary simplifications. After all, it is the utility of most models to simplify while still preserving interesting properties (e.g. somee level of prediction). A major issue is thus validation of models with these, or any other, assumptions by empirical evaluation. Such validation efforts are considered in next section.

5.3 Empirical efforts and results

Finally, considering the found empirical work and how it supports the quantitative methods. The validation variable allows several different methods of validation to support the methods or work under consideration, and how it is distributed is shown in Table 7. Two types of validation are largely dependent on correct formal assumptions: simulation and theoretical approaches. As considered above, these may provide valid results with certain confidence if the underlying assumptions are well-supported by empirical means. The remaining method to provide validated results, for methods and assumptions, is thus empirical which is described below.

Validation	Number
EMP	26
HYP	42
SIM	14
THE	19
None	16
Total	90

Table 7: Validation methods (potentially overlapping except where None) found in work on quantified security. Hypothetical examples seem to dominate effort to validate and present the work. The empirical work is further analyzed in Table 8.

We analyzed the material containing or mentioning empirical work (either by doing evaluation of explicit goals, or describing empirical examples from operational security). Then, it was examined in which way this empirical work is used: *What data was collected, and to which end was it used?*
The result is presented in Table 8, and one can observe that:

- A minority of the surveyed work attempts to empirically evaluate explicit hypotheses with data from multiple sources and environments. The use of data sets is mostly lacking, with the exception of vulnerability models as in [Ozm05], [KMR07], [Ozm07], [AM08].

  First, much of the work falls into either *case studies* of *single systems* with the exception of [Str90], [BAMF01], [MW04], [CH05], [MKW07], [MTMW07], [KDA+07], [AM08].

  Second, much of the work only analyzes systems in operational settings for *limited time* (maximally a few days) with the exception of [JO97, ODK99, KS05a], [LS06], [KDA+07].

- The empirical material varies broadly from collected numerical data to the use of expert judgement in the assessment of quantitative data. This is seen by direct observation of the Data column.

- Experiments have typically not been repeated (been subject to verification or attempts at falsification) using data from different sources. There is one exception: vulnerability discovery models have been compared and examined more than once to different sets of data, with the explicit goal of assessing and predicting vulnerability discovery rates for various software systems[9]. See the Results column.

- Some material mentions empirical results, while keeping details unclear (potentially shrouded behind security classification because of sensitive data). This makes it hard to repeat experiments. See the Data column.

- Several methods are presented with empirical data used to demonstrate *applicability* of methods giving quantitative results [DDK96], [SHJ+02], [HPW03], [MW04], [MBFB05], [MSR07], [MTMW07]. This means showing that a method can produce numbers. However, these efforts do often not attempt validation of result *correctness* (relation to attributes of security targets). Such presentation of numbers as quantitative results does not support the correctness of a method. It is important to distinguish whether a method can generate quantitative outcomes, and to which degree it correctly gives accurate quantitative description of security targets. See the Results column.

5.4 Results from analysis

Based on the findings in previous subsections, we find that a majority of the methods use assumptions that are neither conceptually obvious nor empirically well-tested in operational security. There is even counterevidence in some cases, so work using such assumptions without further validation may be based on faulty assumptions and using methods may give wrong results.
On the other hand, a minority of the work describes empirical efforts which in itself seems initially promising. However, examining the empirical work shows how there is often a lack of validation: efforts have been made at demonstrating how one *applies* methods instead of validating methods to the goal of representing relevant security *attributes* of interest. Vulnerability discovery models are an exception: empirical work have shown the limits of the models proposed so far with respect to predictive power (for which they have been developed). However, models are improving.
In assessing the strength of the quanfication hypothesis, one generally finds a lack of comparison between different methods in using the same kind of methodology or experimenting with same data. There is a lack of empirical evidence that (in the large majority of the cases) either corroborates or falsifies the proposed quantitative methods. Furthermore, almost no solid knowledge exists about relative success between different methods. Little is known about correctness and usability with respect to different security targets

[9]Some vulnerability discovery models have ran into a well-known problem: models using independence have limited success.

	Data	Results
[Str90]	Person survey from 1211 organizations	Data security countermeasures reduce security risk
[DDK96]	3 examples from Unix	Demonstrates MTTF computation, validity unclear
[JO97]	University students, intrusion experiments, one distributed Unix system	Attack phases, effort for nonprofessionals (known vuln.) may be exponentially distributed
[ODK99]	13 known vulnerabilities in large unix system, 2 years, expert weights	3 metrics from different attacker behavior, maybe usable for known vulnerabilities
[Low01]	Red team, 5 days, adversaries weight path risk parameters	Composite metric definition, no computation or validation to goal
[BAMF01]	CERT incident data, self-reporting	Model exploit incident rate, proposed for prediction, demonstrates patching failure
[SHJ+02]	Graph computation, hypothetical network and probabilities	Demonstrates computing automatic attack graphs and probability of intrusion, no validation
[But02]	Interview for security manager risk judgements, survey with a few managers	Cost-benefit approach is popular for management, unclear validation of method
[ADSW03]	Small organizations (20-80 people), experts provide information	Evaluation framework for organizational information security risk, unclear details/quality of result
[HPW03]	A security bulletin (Microsoft), anecdotal evidence	Computation of metric for 3 Windows versions, consistent with chosen/anecdotal evidence
[MW04]	4 versions of Linux operating system, CVE/vulnerability databases	Demonstrating computation on systems, consistent with perceived (subject unknown) beliefs
[Ozm05]	54 months OpenBSD vulnerability data	Reliability models have limited accuracy, null hypothesis for trends not falsified
[MBFB05]	1 SCADA sys, known vulnerabilities, expert judgement	Demonstrating computation of timing metric, unknown validity, consistent with intuition
[KS05a]	20 users, 1 month of virus infections, expert survey and risk judgements	Method gives (internally) consistent results, validity unclear.
[CH05]	Unknown military experts, judgement about 5 categories of 3 systems	Unclear validation of correctness
[KS06]	Attack graph from a network topology, expert opinion, vulnerability database	Demonstrating approach, no validation
[LS06]	Records of port-scans against one university, months data of service interruption	Testing stocahstic model against 4 days of port-scans
[NR06]	VoIP software	Demonstrates method applicable, unclear validation
[MSR07]	3 described (CVE) vulnerabilities	Demonstrates computation with framework, unclear validation to goals
[Ozm07]	Reviewing several modeling attempts	Vulnerability discovery not an independent process
[KMR07]	Web/Database server, Vulnerability discovery data	Vulnerability discovery model for multi-version software, claim to fit
[MKW07]	4 software versions, expert (administrator) survey, security bulletin	Correlation from method attributes to expert judgement, demonstrates application (unvalidated)
[MTMW07]	4 software versions, expert survey, security bulletins, parameter sensitivity	Demonstrates application (result unvalidated)
[KDA+07]	Observed attacks against 35 Internet honeypots, 320 days	Attack time and propagation modeling, attacks probabilistically dependent
[BM07]	Distributed control system at a chemical processing plant	14 metrics, 7 dimensions, demonstrating application, no validation to goal (risk correlation)
[AM08]	4 operating systems, vulnerability databases, 6 vulnerability models	Some vulnerability models are better than others
[GMMS08]	Internet backbone topology (in the US)	Significant difference exists between network vulnerability/performance measures

Table 8: The efforts and results of empirical work. This table summarizes which kind of empirical data was used to which end, in order to assess the degree of support for the different methods. Several methods using empirical work do this in order to produce numbers, without further validation.

in different environments over time: this follows from general lack of comparative and repeated large-sample tests. With respect to the previous terminology in Section 2.1 and the findings above one can currently conclude that *quantified security is a weak hypothesis*.

6. GENERAL DISCUSSION

This section aims to briefly suggest why quantified security currently is a weak hypothesis, what may be done to improve it, and what the risks are with trusting current quantitative models for security assessment and decision-making. The reviewed work has provided many different approaches to quantitative security, but the following issues appear among most of them.

It has obviously been easy to propose a metric or model, but significantly harder to validate that it can be reliably used to describe operational security. Taken together, the field overall has few or no established forms of comparable experimentation which makes it hard to know whether progress is made. This likely depends on the scarce availability and possibility to collect and share such data. This suggests the natural question: how and where could such efforts to validate quantitative methods start? We believe that besides academia there are other actors, such as financial institutions, who have incentives and the mechanisms to start systematic research. These actors increasingly need to acknowledge security risks because of regulations: an example is *operational risk* in Basel II[oIS06], which requires risks with failed processes, people and systems to be quantified.

Furthermore, a number of different perspectives have been applied in the area. Various tools from computer science, reliability theory, economic risk analysis and several other disciplines such as systems management have been proposed to quantify security. While this creates fragmentation between the perspectives, assumptions and targets that are considered as relevant - it is also unclear whether these diverse perspectives can develop common methodology to establish standards in the field. Different security targets have been approached with different perspectives, but it is possible that this phase will be prolonged.

Methodologically, several quantitative methods that have been proposed have the virtue of being simple to apply to obtain some outcome. It has then often been validated that such metrics are *applicable* with respect to counting and data gathering. But often this has lacked validation to their practical *goals* (of representing relevant security attributes). When quantitative methods are proposed without such validation, it is not obvious if this relates to the correctness or the usability of of the methods.

By observing the found empirical material one can observe that work often contains studies particular and single systems or case-studies mostly at a single snapshot in time. Long-term and coordinated collection of data and repeated studies for validity has in many cases not even been mentioned as an important goal for future work. With the different perspectives, this is likely to require developing standards of accurately representing, sharing and storing of data.

This discussion suggests the following ways to improve the knowledge about quantitative methods: compare methods by applying them to the same data (make research that is also problem-driven besides proposing new methods), improve collaboration between different research fields, require better standards of validation, and establish the collection and availability of different data sets. It is outside the scope of this work to suggest how to achieve these means to develop scientifically valid quantified security.

6.1 Problems and risks with lack of validity

Taking a step back, what are the risks from potentially depending on quantitative methods of limited validation? The easy answer is that there is room for research, to corrborate or falsify the methods: but what about using the methods in practice?

For practical application, a significant part of the work in Tables 1, 2 regard improving security decisions. This is often motivated with that security-economic decisions should use quantitative information for rational guidance. However, this survey suggests the risk that such usage may be based on unreliable methods and thus may be irrational for a decision-maker that depends on the methods.

Additionally, there are well-known problems with how quantitative methods may become established in organizations [HK98], and that quantified information [BÖ8, Ver08] may in itself lead to usability problems. An example out of many: with inherent uncertainty about security decisions, people need to justify their decisions [KST02], and this may lead them to overconfidence in available (but unvalidated) quantitative information [KST02].

7. CONCLUSION

This paper surveys 90 papers between 1981 and 2008 to evaluate whether security can be represented using quantitative information (with existing methods). A number of different quantitative methods have been proposed with different perspectives, targets and assumptions. The result from considering a large part of the proposed methods is that quantified security is a weak hypothesis: for most cases, it is unknown if the methods are valid or not in representing operational security.

The foremost reason for this is the lack of repeated large-sample empirical validation of the specific quantiative methods. Explicit empirical validation is usually lacking even where effort with empirical characteristics is found. With a few exceptions, almost no methods have been compared. Another reason is use of model assumptions that lack conceptual or at least empirical support for the targets they are used to model. It is instead possible to find counter evidence to several assumptions.

While a number of theoretical methods have been developed, the availability and the use of solid data seems crucial to allow the field to progress. A study of existing empirical work suggests room for improvement with repeating experiments, the use of shared data sets, and studies that compare different methods.

It should also be apparent that in this paper, we are clearly not attempting to reject modeling or quantification as a fundamentally good idea. However, the effort of the survey allows one to observe limitations in much of the work on quantification so far.

Furthermore, the conclusion depends on natural limits to the effort that was spent: one cannot rule out that there exists unobserved work that is substantially more developed e.g. in terms of descriptive and predictive validation. However, no indication of this was found in the literature that has been considered.

Finally, risks were identified with the use of current and unvalidated quantitative methods, which has recently been advocated by some scholarly work but also by standards organizations. These relate to both economic rationality as well as usability, but also whether we can precisely know if operational security in our systems is getting better or worse. It appears that valid quantification of security is not close but far away on the horizon, and that a number of measures are needed for quantitative security to succeed.

8. ACKNOWLEDGEMENT

The author acknowledges Wolfgang John, Erland Jonsson and Teodor Sommestad for helpful comments. The paper's shepherds Matt Williamson and Rene Rydhof Hansen provided many helpful suggestions and conversations. The NSPW participants helped to create a stimulating environment for discussing these ideas. Support for this work was provided by Swedish Civil Contingencies Agency.

9. REFERENCES

[AAP07] Marco D. Aime, Andrea Atzeni, and Paolo C. Pomi. Ambra: automated model-based risk analysis. In *QoP '07: Proceedings of the 2007 ACM workshop on Quality of protection*, pages 43–48, New York, NY, USA, 2007. ACM.

[AB03] T. Alpcan and T. Basar. A game theoretic approach to decision and analysis in network intrusion detection. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 3, pages 2595–2600 Vol.3, 2003.

[ADSW03] Christopher Alberts, Audrey Dorofee, James Stevens, and Carol Woody. Introduction to the octave approach. Technical report, Carneigie Mellon Software Engineering Institute/US Department of Defense, August 2003.

[AFB95] Jim Alves-Foss and Salvador Barbosa. Assessing computer security vulnerability. *SIGOPS Oper. Syst. Rev.*, 29(3):3–13, July 1995.

[AHP+04] A. Arora, D. Hall, C. A. Piato, D. Ramsey, and R. Telang. Measuring the risk-based value of it security solutions. *IT Professional*, 6(6):35–42, 2004.

[AL05] Andrea Atzeni and Antonio Lioy. Why to adopt a security metric? A brief survey. In *Quality of Protection*, 2005.

[AM07] Ross Anderson and Tyler Moore. The economics of information security: A survey and open questions. In *Fourth bi-annual Conference on the Economics of the Software and Internet Industries*, January 2007.

[AM08] O. H. Alhazmi and Y. K. Malaiya. Application of vulnerability discovery models to major operating systems. *Reliability, IEEE Transactions on*, 57(1):14–22, 2008.

[And01] R. Anderson. Why information security is hard - an economic perspective. In *Computer Security Applications Conference, 2001. ACSAC 2001. Proceedings 17th Annual*, pages 358–365, 2001.

[ANS08] American National Standards Institute (ANSI) / Internet Security Alliance (ISA). *The Financial Impact of Cyber Risk*, 2008.

[Axe00] Stefan Axelsson. The base-rate fallacy and the difficulty of intrusion detection. *ACM Trans. Inf. Syst. Secur.*, 3(3):186–205, 2000.

[BÖ8] Rainer Böhme. Validation of predictions with measurements. In *Dependability Metrics*, pages 14–18. Springer-Verlag, 2008.

[BAMF01] H. K. Browne, W. A. Arbaugh, J. Mchugh, and W. L. Fithen. A trend analysis of exploitations. In *Security and Privacy, 2001. S&P 2001. Proceedings. 2001 IEEE Symposium on*, pages 214–229, 2001.

[BDP07] Stefano Bistarelli, DallaglioMarco, and Pamela Peretti. Strategic games on defense trees. In *Formal Aspects in Security and Trust*, pages 1–15. Springer-Verlag Berlin Heidelberg, 2007.

[BF08] Rainer Böhme and Felix Freiling. On metrics and measurements. In *Dependability Metrics*, pages 7–13. Springer-Verlag, 2008.

[BFP06] Stefano Bistarelli, Fabio Fioravanti, and Pamela Peretti. Defense trees for economic evaluation of security investments. In *ARES '06: Proceedings of the First International Conference on Availability, Reliability and Security*, pages 416–423, Washington, DC, USA, 2006. IEEE Computer Society.

[Bie04] V. Bier. Should the model for security be game theory rather than reliability theory? In *Communications of the Fourth International Conference on Mathematical Methods in Reliability*, 2004.

[Bie06] Vicki Bier. Game-theoretic and reliability methods in counterterrorism and security. In *Statistical Methods in Counterterrorism*, pages 23–40. Springer-Verlag New York, 2006.

[BLP+06] Ahto Buldas, Peeter Laud, Jaan Priisalu, Märt Saarepera, and Jan Willemson. Rational choice of security measures via multi-parameter attack trees. In *Critical Information Infrastructures Security*, pages 235–248. Springer-Verlag Berlin Heidelberg, 2006.

[BM07] Wayne Boyer and Miles Mcqueen. Ideal based cyber security technical metrics for control systems. In *2nd International Workshop on Critical Information Infrastructures Security*, 2007.

[BMG01] Bob Blakley, Ellen Mcdermott, and Dan Geer. Information security is information risk management. In *NSPW '01: Proceedings of the 2001 workshop on New security*

paradigms, pages 97–104, New York, NY, USA, 2001. ACM.

[BOS07] Vicki Bier, Santiago Oliveros, and Larry Samuelson. Choosing what to protect: Strategic defensive allocation against an unknown attacker. *Journal of Public Economic Theory*, 9(4):563–587, August 2007.

[But02] Shawn A. Butler. Security attribute evaluation method: a cost-benefit approach. In *ICSE '02: Proceedings of the 24th International Conference on Software Engineering*, pages 232–240, New York, NY, USA, 2002. ACM.

[CH05] Ping-Teng Chang and Kuo-Chen Hung. Applying the fuzzy-weighted-average approach to evaluate network security systems. *Computers & Mathematics with Applications*, 49(11-12):1797–1814, June 2005.

[CS81] Michael J. Cerullo and Fred A. Shelton. Analyzing the cost-effectiveness of computer controls and security. *The internal auditor*, pages 30–37, October 1981.

[CSS+07] Elizabeth Chew, Marianne Swanson, Kevin Stine, Nadya Bartol, Anthony Brown, and Will Robinson. Nist performance measurement guide for information security (draft). Technical report, NIST, September 2007.

[DDK96] M. Dacier, Y. Deswarte, and M. Kaaniche. Quantitative assessment of operational security: Models and tools, 1996.

[DLK04] R. Dantu, K. Loper, and P. Kolan. Risk management using behavior based attack graphs. In *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on*, volume 1, pages 445–449 Vol.1, 2004.

[DPP05] Zaid Dwaikat and Francesco Parisi-Presicce. Risky trust: risk-based analysis of software systems. *SIGSOFT Softw. Eng. Notes*, 30(4):1–7, July 2005.

[GKMR07] Boaz Golany, Edward H. Kaplan, Abraham Marmur, and Uriel G. Rothblum. Nature plays with dice - terrorists do not: Allocating resources to counter strategic versus probabilistic risks. *European Journal of Operational Research*, In Press, Corrected Proof, 2007.

[GMMS08] Tony H. Grubesic, Timothy C. Matisziw, Alan T. Murray, and Diane Snediker. Comparative approaches for assessing network vulnerability. *International Regional Science Review*, 31(1):88–112, January 2008.

[GMT05] Christopher Griffin, Bharat Madan, and Kishor Trivedi. State space approach to security quantification. In *COMPSAC '05: Proceedings of the 29th Annual International Computer Software and Applications Conference (COMPSAC'05) Volume 2*, pages 83–88, Washington, DC, USA, 2005. IEEE Computer Society.

[Hau07] Kjell Hausken. Protecting complex infrastructures against strategic attackers. Technical report, Faculty of Social Sciences, University of Stavanger, 2007.

[HHH06] Jonas Hallberg, Niklas Hallberg, and Amund Hunstad. Crossroads and XMASS: Framework and method for system it security assessment. Technical report, FOI, Swedish Defence Research Agency, 2006.

[HK98] John Hauser and Gerald Katz. Metrics: you are what you measure! *European Management Journal*, 16(5):517–528, October 1998.

[HMOS02] S. N. Hamilton, W. N. Miller, A. Ott, and O. S. Saydjari. The role of game theory in information warfare. In *4th Information survivability workshop, (ISW-2001/2002)*, 2002.

[HPW03] M. Howard, J. Pincus, and J. M. Wing. Measuring relative attack surfaces. In *Proc. of Workshop on Advanced Developments in Software and Systems Security*, 2003.

[HSHJ08] T. Heyman, R. Scandariato, C. Huygens, and W. Joosen. Using security patterns to combine security metrics. In *Availability, Reliability and Security, 2008. ARES 08. Third International Conference on*, pages 1156–1163, 2008.

[Hul08] Rolf Hulthén. Communicating the economic value of security investments; value at security risk. In *Workshop on the Economics of Information Security*, 2008.

[JO97] Erland Jonsson and Tomas Olovsson. A quantitative model of the security intrusion process based on attacker behavior. *IEEE Transactions on Software Engineering, Vol. 23, No. 4, April*, 1997.

[JW07] Aivo Jürgenson and Jan Willemson. Processing multi-parameter attacktrees with estimated parameter values. In *Advances in Information and Computer Security*, pages 308–319. Springer-Verlag Berlin Heidelberg, 2007.

[Kah00] Daniel Kahneman. *Choices, Values, and Frames*. Cambridge University Press, September 2000.

[KDA+07] Mohamed Kaaniche, Y. Deswarte, Eric Alata, Marc Dacier, and Vincent Nicomette. Empirical analysis and statistical modeling of attack processes based on honeypots, Apr 2007.

[KH03] Howard Kunreuther and Geoffrey Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2):231–249, March 2003.

[KMR07] Jinyoo Kim, Yashwant K. Malaiya, and Indrakshi Ray. Vulnerability discovery in multi-version software systems. In *High Assurance Systems Engineering Symposium, 2007. HASE '07. 10th IEEE*, pages 141–148, 2007.

[KS05a] Bilge Karabacak and Ibrahim Sogukpinar. ISRAM: information security risk analysis method. *Computers and Security*, 24(2):147 – 159, 2005.

[KS05b] Igor Kotenko and Mihail Stepashkin. Analyzing vulnerabilities and measuring security level at design and exploitation stages of computer network life cycle. In *Computer Network Security*, pages 311–324. Springer-Verlag Berlin Heidelberg, 2005.

[KS06] Igor Kotenko and Mikhail Stepashkin. Attack graph based evaluation of network security. In *Communications and Multimedia Security*, pages 216–227. Springer-Verlag Berlin Heidelberg, 2006.

[KST82] Daniel Kahneman, Paul Slovic, and Amos Tversky. *Judgment under Uncertainty : Heuristics and Biases*. Cambridge University Press, April 1982.

[KST02] Daniel Kahneman, Paul Slovic, and Amos Tversky. *Heuristics and Biases: The psychology of intuitive judgement*. Cambridge University Press, 2002.

[LBF+93] B. Littlewood, S. Brocklehurst, N. Fenton, P. Mellor, S. Page, D. Wright, J. Dobson, J. Mcdermid, and D. Gollmann. Towards operational measures of computer security. *Journal of Computer Security*, 2:211–229, 1993.

[LJ08] D. J. Leversage and E. James. Estimating a system's mean time-to-compromise. *Security & Privacy, IEEE*, 6(1):52–60, 2008.

[Low01] John Lowry. An initial foray into understanding adversary planning and courses of action. *DARPA Information Survivability Conference and Exposition,*, 1:0123, 2001.

[LS06] Vincent C. S. Lee and Linyi Shao. Estimating potential it security losses: An alternative quantitative approach. *Security & Privacy, IEEE*, 4(6):44–52, 2006.

[LW05] Kong-Wei Lye and Jeannette M. Wing. Game strategies in network security. *International Journal of Information Security*, 4(1):71–86, February 2005.

[LZY05] Peng Liu, Wanyu Zang, and Meng Yu. Incentive-based modeling and inference of attacker intent, objectives, and strategies. *ACM Trans. Inf. Syst. Secur.*, 8(1):78–118, February 2005.

[MBFB05] M. A. Mcqueen, W. F. Boyer, M. A. Flynn, and G. A. Beitel. Time-to-compromise model for cyber risk reduction estimation. In *Quality of Protection*, 2005.

[MBFB06] Miles A. Mcqueen, Wayne F. Boyer, Mark A. Flynn, and George A. Beitel. Quantitative cyber risk reduction estimation methodology for a small SCADA control system. In *HICSS '06: Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, Washington, DC, USA, 2006. IEEE Computer Society.

[Mcd05] J. Mcdermott. Attack-potential-based survivability modeling for high-consequence systems. In *Information Assurance, 2005. Proceedings. Third IEEE International Workshop on*, pages 119–130, 2005.

[Mer99] James W. Meritt. A method for quantitative risk analysis. In *Proceedings of the 22nd National Information Systems*

Security Conference, 1999.

[MGPVT02] B. B. Madan, K. Gogeva-Popstojanova, K. Vaidyanathan, and K. S. Trivedi. Modeling and quantification of security attributes of software systems. In *Proceedings of the International Conference on Dependable Systems and Networks*, pages 505–514, 2002.

[MGPVT04] Bharat B. Madan, Katerina Goseva-Popstojanova, Kalyanaraman Vaidyanathan, and Kishor S. Trivedi. A method for modeling and quantifying the security attributes of intrusion tolerant systems. *Perform. Eval.*, 56(1-4):167–186, 2004.

[MKF03] J. Mcdermott, A. Kim, and J. Froscher. Merging paradigms of survivability and security: stochastic faults and designed faults. In *NSPW '03: Proceedings of the 2003 workshop on New security paradigms*, pages 19–25, New York, NY, USA, 2003. ACM.

[MKW07] P. K. Manadhata, D. K. Kaynar, and J. M. Wing. A formal model for a systems attack surface. Technical report, Carnegie Mellon University, 2007.

[MSR07] Peter Mell, Karen Scarfone, and Sasha Romanosky. *CVSS: A Complete Guide to the Common Vulnerability Scoring Systems Version 2.0*. FIRST: Forum of Incident Response and Security Teams, June 2007.

[MT04] B. B. Madan and K. S. Trivedi. Security modeling and quantification of intrusion tolerant systems using attack-response graph. *J. High Speed Netw.*, 13(4):297–308, October 2004.

[MTMW07] P. K. Manadhata, K. M. C. Tan, R. A. Maxion, and J. M. Wing. An approach to measuring a systems attack surface. Technical report, School of Computer Science, Carnegie Mellon University, 2007.

[MW04] P. Manadhata and J. M. Wing. Measuring a system's attack surface. Technical report, Carnegie Mellon University, 2004.

[MW05] P. Manadhata and J. Wing. An attack surface metric. Technical report, Carnegie Mellon University, 2005.

[MYY+07] Dapeng Man, Wu Yang, Yongtian Yang, Wei Wang, and Lejun Zhang. A quantitative evaluation model for network security. In *Computational Intelligence and Security, 2007 International Conference on*, pages 773–777, 2007.

[Nic05] D. M. Nicol. Modeling and simulation in security evaluation. *Security & Privacy, IEEE*, 3(5):71–74, 2005.

[NR06] Syed Naqvi and Michel Riguidel. Quantifiable security metrics for large scale heterogeneous systems. In *Carnahan Conferences Security Technology, Proceedings 2006 40th Annual IEEE International*, pages 209–215, 2006.

[NST04] D. M. Nicol, W. H. Sanders, and K. S. Trivedi. Model-based evaluation: from dependability to security. *Dependable and Secure Computing, IEEE Transactions on*, 1(1):48–65, 2004.

[ODK99] Rodolphe Ortalo, Yves Deswarte, and Mohamed Kaâniche. Experimenting with quantitative evaluation tools for monitoring operational security. *IEEE Trans. Softw. Eng.*, 25(5):633–650, September 1999.

[oIS06] Bank of International Settlements. Basel II: International convergence of capital measurement and capital standards: a revised framework. Online publication, June 2006.

[Ozm05] Andy Ozment. Software security growth modeling: Examining vulnerabilities with reliability growth models. In *Quality of Protection*, 2005.

[Ozm07] Andy Ozment. Improving vulnerability discovery models. In *QoP '07: Proceedings of the 2007 ACM workshop on Quality of protection*, pages 6–11, New York, NY, USA, 2007. ACM.

[Pay06] S. C. Payne. A guide to security metrics. Technical report, SANS Institute, 2006.

[PJAS06] Joseph Pamula, Sushil Jajodia, Paul Ammann, and Vipin Swarup. A weakest-adversary security metric for network configuration security analysis. In *QoP '06: Proceedings of the 2nd ACM workshop on Quality of protection*, pages 31–38, New York, NY, USA, 2006. ACM.

[Pop59] Karl R. Popper. *The Logic of Scientific Discovery*. Springer, 1959.

[PPN06] Victor-Valeriu Patriciu, Iustin Priescu, and Sebastian Nicolaescu. Security metrics for enterprise information systems. *Journal of Applied Quantitative Methods*, pages 151–159, 2006.

[RGH07] P. A. S. Ralston, J. H. Graham, and J. L. Hieb. Cyber security risk assessment for SCADA and DCS networks. *ISA Transactions*, 46(4):583–594, October 2007.

[SBS+03] Marianne Swanson, Nadya Bartol, John Sabato, Joan Hash, and Laurie Graffo. Security metrics guide for information technology systems. Technical report, NIST, 2003.

[Sch99] Bruce Schneier. Attack trees. *Dr. Dobb's Journal*, 1999.

[Sch02] Stuart Schechter. Quantitatively differentiating system security. In *Workshop on the Economics of Information Security*, 2002.

[Sch04] S. E. Schechter. Toward econometric models of the security risk from remote attacks. *Security & Privacy, IEEE*, 3(1):40–44, 2004.

[Sch07] Bruce Schneier. The psychology of security, 2007.

[SCHB07] Dan Shen, Genshe Chen, Leonard Haynes, and Erik Blasch. Strategies comparison for game theoretic cyber situational awareness and impact assessment. In *Information Fusion, 2007 10th International Conference on*, pages 1–8, 2007.

[SdBF+02] Ketil Stolen, Folker den Braber, Rune Fredriken, Bjorn Axel Gran, Siv-Hilde Houmb, Mass Soldal Lund, Yahhis C. Stamatio, and Jan Oyvind Aagedal. Model-based risk assessment - the coras approach. In *Proc. Norsk Informatikkkonferanse (NIK'2002)*, pages 239–249, 2002.

[SGF02] Gary Stoneburner, Alice Goguen, and Alexis Feringa. Risk management guide for information technology systems. Technical report, Information Technology Laboratory, National Institute of Standards and Technology, 2002.

[SH00] Kevin J. Soo Hoo. How Much Is Enough? A Risk-Management Approach to Computer Security. Technical report, Consortium for Research on Information Security and Policy (CRISP), 2000.

[SH02] Kevin J. Soo Hoo. How Much Is Enough? A Risk Management Approach to Computer Security. In *Workshop on the Economics of Information Security*, 2002.

[SH03] Bomil Suh and Ingoo Han. The is risk analysis based on a business model. *Inf. Manage.*, 41(2):149–158, 2003.

[SHJ+02] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. M. Wing. Automated generation and analysis of attack graphs. In *Proceedings of 2002 IEEE Symposium on Security and Privacy*, pages 273–284, 2002.

[SHK06] Karin Sallhammar, Bjarne E. Helvik, and Sven J. Knapskog. A game-theoretic approach to stochastic security and dependability evaluation. In *Dependable, Autonomic and Secure Computing, 2nd IEEE International Symposium on*, pages 61–68, 2006.

[SHK07] Karin Sallhammar, Bjarne E. Helvik, and Svein J. Knapskog. A framework for predicting security and dependability measures in real-time. *International Journal of Computer Science and Network Security*, 7(3):169–183, 2007.

[Sim55] Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118, 1955.

[SKH05] K. Sallhammar, S. J. Knapskog, and B. E. Helvik. Using stochastic game theory to compute the expected behavior of attackers. In *Applications and the Internet Workshops, 2005. Saint Workshops 2005. The 2005 Symposium on*, pages 102–105, 2005.

[SLN04] Sankalp Singh, James Lyons, and David M. Nicol. Fast model-based penetration testing. In *WSC '04: Proceedings of the 36th conference on Winter simulation*, pages 309–317. Winter Simulation Conference, 2004.

[SPG98] Laura Painton Swiler, Cynthia Philips, and Philips Gaylor. A graph-based network-vulnerability analysis system. Technical report, SANDIA, 1998.

[Str90] Detmar W. Straub. Effective is security: An empirical study. *Information Systems research*, 1(3):255–276, September 1990.

[SW00] Gregg Schudel and Bradley Wood. Adversary work factor as a metric for information assurance. In *NSPW '00: Proceedings of the 2000 workshop on New security paradigms*, pages 23–30, New York, NY, USA, 2000. ACM.

[Ver08] Vilhelm Verendel. A prospect theory approach to security. Technical report, Department of Computer Science and Engineering, Chalmers University of Technology, 2008.

[VGM+96] J. Voas, A. Ghosh, G. Mcgraw, F. Charron, and K. Miller. Defining an adaptive software security metric from a dynamic software failure tolerance measure. In *Computer Assurance, 1996. COMPASS '96, 'Systems Integrity. Software Safety. Process Security'. Proceedings of the Eleventh Annual Conference on*, pages 250–263, 1996.

[VMP04] Carlos Villarrubia, Eduardo F. Medina, and Mario Piattini. Towards a classification of security metrics. In *WOSIS*, pages 342–350, 2004.

[Waw06] Dariusz Wawrzyniak. Information security risk assessment model for risk management. *Trust and Privacy in Digital Business*, pages 21–30, 2006.

[WIL+08] Lingyu Wang, Tania Islam, Tao Long, Anoop Singhal, and Sushil Jajodia. An attack graph-based probabilistic security metric. In *Proceeedings of the 22nd annual IFIP WG 11.3 working conference on Data and Applications Security*, pages 283–296. Springer-Verlag Berlin Heidelberg, 2008.

[WSJ07] Lingyu Wang, Anoop Singhal, and Sushil Jajodia. Toward measuring network security using attack graphs. In *QoP '07: Proceedings of the 2007 ACM workshop on Quality of protection*, pages 49–54, New York, NY, USA, 2007. ACM.

[WT06] Max Walter and Carsten Trinitis. Quantifying the security of composed systems. *Parallel Processing and Applied Mathematics*, pages 1026–1033, 2006.

[WW97] C. Wang and W. Wulf. Towards a framework for security measurement. In *NISSC*, 1997.

[YCL06] Fu-Hong Yang, Chi-Hung Chi, and Lin Liu. A risk assessment model for enterprise network security. In *Autonomic and Trusted Computing*, pages 293–301. Springer-Verlag Berlin Heidelberg, 2006.

[YSH+08] A. Yautsiukhin, R. Scandariato, T. Heyman, F. Massacci, and W. Joosen. Towards a quantitative assessment of security in software architectures. In *Proceedings of the 13th Nordic Workshop on Secure IT Systems*, 2008.

[ZWW06] Guosheng Zhao, Huiqiang Wang, and Jian Wang. A novel quantitative analysis method for network survivability. In *Computer and Computational Sciences, 2006. IMSCCS '06. First International Multi-Symposiums on*, volume 2, pages 30–33, 2006.