When Not All Bits Are Equal: Worth-Based Information Flow

Mário S. Alvim[1], Andre Scedrov[2], and Fred B. Schneider[3]

[1] Universidade Federal de Minas Gerais, Brazil
[2] University of Pennsylvania, USA
[3] Cornell University, USA

Abstract. Only recently have approaches to quantitative information flow started to challenge the presumption that all leaks involving a given number of bits are equally harmful. This paper proposes a framework to capture the semantics of information, making quantification of leakage independent of the syntactic representation of secrets. Secrets are defined in terms of fields, which are combined to form structures; and a *worth assignment* is introduced to associate each structure with a worth (perhaps in proportion to the harm that would result from disclosure). We show how worth assignments can capture inter-dependence among structures within a secret, modeling: (i) secret sharing, (ii) information-theoretic predictors, and (iii) computational (as opposed to information-theoretic) guarantees for security. Using non-trivial worth assignments, we generalize Shannon entropy, guessing entropy, and probability of guessing. For deterministic systems, we give a lattice of information to provide an underlying algebraic structure for the composition of attacks. Finally, we outline a design technique to capture into worth assignments relevant aspects of a scenario of interest.

1 Introduction

Quantitative information flow (QIF) is concerned with measuring how much secret information leaks to an adversary through a system. The adversary is presumed to have *a priori information* about the secrets before execution starts and to access *public observables* as execution proceeds. By combining a priori information and public observables, the adversary achieves *a posteriori information* about the secrets. The *leakage* from an execution is the difference between a posteriori and a priori information.

This definition of leakage depends on how information is measured. Cachin [1] advocates that information measures not only include a way to calculate some numeric value but also offer an *operational interpretation*, which describes what aspect of interest is being quantified. Popular information measures include: *Shannon entropy* [2–8], which measures how much information is leaked per guess; *guessing entropy* [9, 10], which measures how many tries are required before the secret is correctly guessed; and *probability of guessing* [11, 12], which measures how likely it is that a secret is correctly inferred in a certain number of tries.

M. Abadi and S. Kremer (Eds.): POST 2014, LNCS 8414, pp. 120–139, 2014.

These measures are best suited to sets of monolithic and equally valuable secrets, so researchers have recently begun to consider richer scenarios. The *g-leakage* framework [13] of Alvim et al. makes use of gain functions to quantify the benefit of different guesses for the secret. However, identifying sufficiently expressive yet not over-complicated gain-functions is often a challenge. Moreover, that framework generalizes probability of guessing, but not Shannon entropy or guessing entropy. Finally, it is not suitable to infinitely risk-averse adversaries. In this paper we propose an approach that addresses these limitations; a detailed comparison with g-leakage is given in Section 6.

We model a secret as being partitioned into *fields*, which are combined to form *structures*. Since disclosure of different structures might cause different harms, a *worth assignment* is introduced to associate a *worth* with each structure. For instance, the secret corresponding to a client's bank account might comprise two 10-digit structures: a pincode and a telephone number. Leaking the pincode has the potential to cause considerable harm, so that structure would be assigned high worth; the telephone number is public information, so this structure would be assigned low worth.

Assuming that all structures have equal worth can lead to misleading comparisons between systems that leak structures with different worths but the same numbers of bits. Conversely, ignoring the structure of secrets may lead to a deceptive estimate of the harm from leaking different numbers of bits. Consider two systems that differ in the way they represent a house address. In system C_1, standard postal addresses are used (i.e., a number, street name, and zip-code); system C_2 uses GPS coordinates (i.e., a latitude and a longtitude, each a signed 10-digit number). Under Shannon entropy with plausible sizes[1] for address fields, C_1 requires 129 bits to represent a location that C_2 represents using 69 bits. Yet the same content is revealed whether C_1 leaks its 129 bits or C_2 leaks its 69 bits. (The a priori information for addresses in C_1 is not zero, since certain values for a house number, street name, and zip-code can be ruled out. And a similar argument can be made for C_2, given knowledge of habitable terrain. Accounting for idiosyncrasies in the syntactic representation of secrets, however, can be a complicated task, hence an opportunity for error. Worth assignments avoid some of that complexity.)

When secrets are not modeled as monolithic, distinct structures within a given secret may be correlated. A clever adversary, thus, might infer information about a structure with more worth (and presumably better protected) by attacking a correlated structure with less worth (and presumably less well protected). For instance, the location of a neighborhood is often correlated to the political preferences of its residents, so an adversary may target a person's house address to infer information about what political party they support. Worth assignments can model such correlations and adjust the relative worth of structures. Moreover, they can capture the computational complexity of inferring one structure from the other, which is a common limitation of information theoretical

[1] Specifically, assume a 5-digit house number, a 20-character alphabetic street name, and a 5-digit zip-code.

approaches to QIF. As an example, a public RSA key is a perfect predictor, in an information theoretical sense, for the corresponding private key. In practice, however, the public key should not be assigned the same worth as the private key because a realistic adversary is not expected to retrieve the latter from the former in viable time.

In this paper, we propose *measures of information worth* that incorporate the structure and worth of secrets. As in other QIF literature, we assume the adversary performs attacks, controlling the low input to a probabilistic system execution and observing the low outputs. An attack induces a probability distribution on the space of secrets according to what the adversary observes. This characterization admits measures of information worth for the information contained in each distribution; leakage is then defined as the difference in information between distributions. Our approach generalizes probability of guessing, guessing entropy, and Shannon entropy to admit non-trivial worth assignments. Yet our work remains consistent with the Lattice of Information [14] for deterministic systems, which is an underlying algebraic structure for sets of system executions.

The main contributions of this paper are:

- We propose a framework of structures and worth assignments to capture the semantics of information, making the quantification of leakage independent of the particular representation chosen for secrets.
- We show how to use worth assignments to model the inter-dependence among structures within a given secret, capturing practical scenarios including: (i) secret sharing, (ii) information-theoretic predictors, and (iii) computational (as opposed to information-theoretic) guarantees for security.
- We generalize Shannon entropy and guessing entropy to incorporate worth explicitly, and we introduce other measures without traditional equivalents. We show that our theory of measures of information worth and the g-leakage framework are not comparable in general, although they do overlap.
- We prove that our measures of information worth are consistent with the Lattice of Information for deterministic systems, which allows sound reasoning about the composition of attacks in such systems.
- We outline a design technique for worth assignments that capture the following aspects of the scenario of interest: (i) *secrecy requirements* that determine what structures are intrinsically sensitive, and by how much, (ii) *consistency requirements* that ensure the adequacy of the worth assignment, and (iii) the *adversarial knowledge* that may be of help in attacks.

The paper is organized as follows. Section 2 describes our model for the structure and worth of secrets in probabilistic systems. Section 3 uses worth assignments to propose measures of information worth. Section 4 shows that the proposed measures are consistent with respect to the Lattice of Information for deterministic systems under composite attacks. Section 5 outlines a technique for designing adequate worth assignments for a scenario of interest. Finally, Section 6 discusses related work, and Section 7 concludes the paper. Full proofs can be found in the Appendix of the corresponding technical report [15].

2 Modeling the Structure and Worth of Secrets

We decompose secrets into elementary units called *fields*, each a piece of information with a domain. Let $\mathcal{F} = \{f_1, \dots, f_m\}$ denote the (finite) set of fields in some scenario of interest, and for $1 \leq i \leq m$, let $domain(f_i)$ be the domain of values for field f_i. A *structure* is a subset $\mathfrak{f} \subseteq \mathcal{F}$, and if $\mathfrak{f} = \{f_{i_1}, \cdots, f_{i_k}\}$, its domain is given by $domain(\mathfrak{f}) = domain(f_{i_1}) \times \cdots \times domain(f_{i_k})$. The set of all possible structures is the power set $\mathcal{P}(\mathcal{F})$ of fields, and the structure $\mathfrak{f} = \mathcal{F}$ containing all fields is called the *maximal structure*.

A *secret* s is a mapping from the maximal structure to values, i.e., $s = \langle s[f_1], \dots, s[f_m] \rangle$, where $s[f_i] \in domain(f_i)$ is the value assumed by field f_i. Hence the set $\mathcal{S}$ of possible secrets is $\mathcal{S} = domain(\mathcal{F})$. Given a secret s and a (not necessarily maximal) structure $\mathfrak{f} \subseteq \mathcal{F}$, we call a *sub-secret* $s[\mathfrak{f}]$ the projection of s on the domain of $\mathfrak{f}$, and the set of all possible sub-secrets associated with that structure is $\mathcal{S}[\mathfrak{f}] = domain(\mathfrak{f})$.

Structures may carry some valuable piece of information on their own. A *worth assignment* attributes to each structure a non-negative, real number. Worth may be seen as the utility obtained by an adversary who learns the contents of the structure, or it may be seen as the damage suffered should the contents of the structure become known to that adversary.

Definition 1 (Worth assignment). *A* worth assignment *is a function* $\omega : \mathcal{P}(\mathcal{F}) \to \mathbb{R}$ *from the set of structures to reals, satisfying for all* $\mathfrak{f}, \mathfrak{f}' \in \mathcal{P}(\mathcal{F})$*:* (i) non-negativity*:* $\omega(\mathfrak{f}) \geq 0$*, and* (ii) monotonicity*:* $\mathfrak{f} \subseteq \mathfrak{f}' \implies \omega(\mathfrak{f}) \leq \omega(\mathfrak{f}')$*.*

We require non-negativity of ω because the knowledge of the contents of a structure should not carry a negative amount of information, and we require monotonicity because every structure should be at least as sensitive as any of its parts. Note that monotonicity implies that the worth of the maximal structure, $\omega(\mathcal{F})$, is an upper bound for the worth of every structure.

Expressiveness of Worth Assignments. The worth of a structure should appropriately represent the sensitivity of that structure in a scenario of interest. Consider a medical database where a secret is a patient's entire record, and structures are sub-sets of that record (e.g., a patient's name, age, smoking habits). The worth assigned to an individual's smoking habits should reflect: (i) how much the *protector* (i.e., the party interested in keeping the secret concealed) cares about hiding whether an individual is a smoker, (ii) how much an adversary would benefit from learning whether an individual is a smoker, and, more subtly, (iii) how effective (information-theoretically and/or computationally) a predictor an individual's smoking habits are for other sensitive structures (for instance, heavy smokers are more likely to develop lung cancer, and insurance companies may deny them coverage based on that). Worth assignments can capture these aspects, modeling also:

a) **Semantic-based leakage.** Worth assignments provide a natural means to abstract from syntactic idiosyncrasies and treat structures according to meaning. In the bank system of Section 1, for instance, we would assign higher worth to the 10-digit pincode than to the 10-digit telephone number, thus distinguishing among eventual 10-digit leaks according to relevance:

$$\omega(\{\textsf{pin-code}\}) > \omega(\{\textsf{telephone number}\}).$$

Conversely, structures with equivalent meanings should be assigned the same worth, regardless of representation. For instance, the worth of all structures corresponding to an address should be the same, whether it is represented in GPS coordinates or in the standard postal address format:

$$\omega(\{\textsf{GPS address}\}) = \omega(\{\textsf{postal address}\}).$$

b) **Secret sharing.** The combination of two structures may convey more worth than the sum of their individual worths. In *secret sharing*, for instance, different persons retain distinct partial secrets (i.e., structures) that in isolation give no information about the secret as a whole (i.e., the maximal structure), but that reveal the entire secret when combined. As another example, a decryption key without any accompanying ciphertext is of little worth, so each corresponding structure should have, in isolation, a worth close to zero. When combined, however, the benefit to the adversary exceeds the sum of their individual worths:

$$\omega(\{\textsf{ciphertext}, \textsf{decryption key}\}) \gg \omega(\{\textsf{ciphertext}\}) + \omega(\{\textsf{decryption key}\}).$$

c) **Correlation of structures.** Knowledge of a particular structure may imply knowledge of another (e.g., if the adversary has access to tax files, learning someone's tax identification number implies learning their name as well), or it may increase the probability of learning another structure (recall the correlation between smoking habits and lung cancer). An adversary might exploit correlations between different structures within a given secret to obtain information about a more important (and presumably better protected) structure through a less important (and presumably less well protected) structure. By considering the distribution on secrets and the capabilities of the adversary, we can adjust the relative worth of one structure with respect to any other, thus avoiding potentially harmful loopholes. In particular, worth assignments can model:

   (i) **Information-theoretic predictors.** The worth of a structure should reflect the worth it carries, via correlation, from other structures. For instance, when an individual's identity can be recovered with 60% probability from the combination of the zip-code, date of birth, and gender [16], we might enforce $\omega(\{\textsf{zip-code}, \textsf{date of birth}, \textsf{gender}\})$ to be at least as great as 60% of the worth $\omega(\{\textsf{identity}\})$. More generally, given any two structures $\mathfrak{f}, \mathfrak{f}' \in \mathcal{P}(\mathcal{F})$, the requirement

$$\omega(\mathfrak{f}) \geq \textit{correlation}(\mathfrak{f}, \mathfrak{f}') \cdot \omega(\mathfrak{f}')$$

   might be imposed on a worth assignment ω. Here $\textit{correlation}(\mathfrak{f}, \mathfrak{f}')$ is a function representing how well $\mathfrak{f}$ predicts $\mathfrak{f}'$.

(ii) **Computational effort.** Even perfect information-theoretic correlations among structures may not be of practical use for the adversary (e.g., the correlation of public and private RSA keys). Worth assignments can reflect this. We can impose, on any two structures $\mathfrak{f}, \mathfrak{f}' \in \mathcal{P}(\mathcal{F})$, the requirement

$$\omega(\mathfrak{f}) > \omega(\mathfrak{f}')/\mathit{cost}(\mathfrak{f}, \mathfrak{f}'),$$

where $\mathit{cost}(\mathfrak{f}, \mathfrak{f}')$ is a function of the computational effort needed to obtain $\mathfrak{f}'$ from $\mathfrak{f}$.

2.1 A Worth-Based Approach to QIF

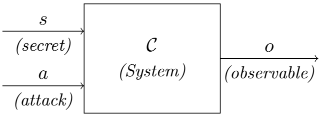

Fig. 1. A system with one high input, one low input, and one low output

We adopt a probabilistic version of the model of deterministic systems and attacks proposed by Köpf and Basin [17]. Let $\mathcal{S}$ be a finite set of *secrets*, $\mathcal{A}$ be a finite set of adversary-controlled inputs or *attacks*, and $\mathcal{O}$ be a finite set of *observables*. A *(probabilistic computational) system* is a family $C = \{(\mathcal{S}, \mathcal{O}, C_a)\}_{a \in \mathcal{A}}$ of *(information-theoretic) channels* parametrized by the adversary-chosen input $a \in \mathcal{A}$. Each $(\mathcal{S}, \mathcal{O}, C_a)$ is a channel in which $\mathcal{S}$ is the *channel input*, $\mathcal{O}$ is the *channel output*, and C_a is a $|\mathcal{S}| \times |\mathcal{O}|$ matrix of conditional probability distributions called the *channel matrix*. Each entry $C_a(s, o)$ in the matrix represents the probability of the system producing observable o when the secret is s and the adversary-chosen low input is a. Given a probability distribution $p_\mathcal{S}$ on $\mathcal{S}$, the behavior of the system under attack a is described by the joint distribution $p_a(s, o) = p_\mathcal{S}(s) \cdot C_a(s, o)$, with marginal $p_a(o) = \sum_s p_a(s, o)$, and conditional distribution $p_a(s|o) = p_a(s, o)/p_a(o)$ whenever $p_a(o) > 0$ (and similarly for $p_a(s)$ and $p_a(o|s)$).

As is usual in QIF, assume that the adversary knows the probability distribution $p_\mathcal{S}$ on the set of secrets and the family of channel matrices C describing the system's behavior. By controlling the low input, the adversary can launch an attack as follows: pick $a \in \mathcal{A}$ so the channel matrix is set to C_a, thereby manipulating the behavior of the system. The adversary's goal is to infer as much information as possible from the secret, given knowledge about how the system works, the attack fed to the system, and the observations made as the system executes.

Let Ω be the set of all possible worth assignments for the structures of $\mathcal{S}$, $Pr(S)$ be the set of all probability distributions on $\mathcal{S}$, and $\mathcal{C}_\mathcal{A}$ be the set of channel matrices induced by attacks $a \in \mathcal{A}$. A *measure of information worth* is a function $\nu : \Omega \times Pr(\mathcal{S}) \times \mathcal{C}_\mathcal{A} \to \mathbb{R}^+$. The quantity $\nu(\omega, p_\mathcal{S}, C_a)$ represents the *a posteriori* information with respect to S revealed by attack $C_a \in \mathcal{C}_\mathcal{A}$, given probability distribution $p_\mathcal{S} \in Pr(\mathcal{S})$ on secrets and worth assignment $\omega \in \Omega$. Before any attack is performed, the adversary has some *a priori* information about the secret due to knowledge of $p_\mathcal{S}$ and ω only, and we represent this information by

$\nu(\omega, p_S)$. Because the attack is expected to disclose secret information to the adversary, the leakage from an attack C_a is defined as the difference[2], between the a posteriori and a priori information associated with C_a.

Before discussing measures of information, we will fix some additional notation. For any $\mathcal{S}' \subseteq \mathcal{S}$ we denote by $p_S(\cdot|\mathcal{S}')$ the normalization of p_S with respect to $\mathcal{S}'$, i.e., for every $s \in \mathcal{S}$, $p_S(s|\mathcal{S}') = p_S(s)/p_S(\mathcal{S}')$ if $s \in \mathcal{S}'$, and $p_S(s|\mathcal{S}') = 0$ otherwise. The support of a distribution p_S is denoted $supp(p_S)$. A set $\mathsf{P} = \{\mathcal{S}_1, \ldots, \mathcal{S}_n\}$ is a *partition* on $\mathcal{S}$ iff: (i) $\bigcup_{\mathcal{S}_i \in \mathsf{P}} \mathcal{S}_i = \mathcal{S}$, and (ii) for $1 \leq i \neq j \leq n$, $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset$. Each $\mathcal{S}_i \in \mathsf{P}$ is called a *block* in the partition. We denote the set of all partitions in $\mathcal{S}$ by $\texttt{LoI}(\mathcal{S})$ [3]. Following [10], any partition $\mathsf{P}_a = \{\mathcal{S}_{o_1}, \ldots, \mathcal{S}_{o_n}\}$ on $\mathcal{S}$ induced by the attack a can be seen as a random variable with carrier $\{\mathcal{S}_{o_1}, \ldots, \mathcal{S}_{o_n}\}$ and probability distribution $p_S(\mathcal{S}_{o_i}) = \sum_{s \in \mathcal{S}_{o_i}} p_S(s)$.

3 Measures of Information Worth

3.1 Operational Interpretation of Measures Revisited

One of Shannon's greatest insights, which ultimately led to the creation of the field of information theory, can be formulated as: *information is describable in terms of answers to questions.* The more information the adversary has about a random variable, the fewer questions of a certain type that must be asked in order to infer its value, and the smaller the Shannon entropy of this random variable.

Formally, the *Shannon entropy* of a probability distribution p_S is defined as $SE(p_S) = -\sum_s p_S(s) \log p_S(s)$, and the *conditional Shannon entropy* of p_S given a channel C_a is defined as $SE(p_S, C_a) = \sum_{o \in \mathcal{O}} p_a(o) SE(p_a(\cdot|o))$. A possible operational interpretation of this measure is: The adversary can pose questions *Does* $S \in \mathcal{S}'$*?*, for some $\mathcal{S}' \subseteq \mathcal{S}$, to an oracle, and Shannon entropy quantifies the expected minimum number of guesses needed to infer the entire secret with certainty. A decrease in the Shannon entropy of the secret space caused by a system can be seen as the leakage from the system. This question-and-answer interpretation has an algorithmic equivalent: $\mathcal{S}$ is seen as a search space, and by repeatedly asking questions *Does* $S \in \mathcal{S}'$*?*, the adversary is performing a binary search on the space of secrets. Now, Shannon entropy corresponds to the average height of the optimal binary search tree.

However, Shannon entropy is not the unique meaningful measure of information. Guessing entropy allows the adversary to pose a different type of question; whereas probability of guessing quantifies a different aspect of the scenario of

[2] Braun et al. [12] make a distinction between this definition of leakage, called *additive leakage*, and *multiplicative leakage*, where the ratio (rather than the difference) of the a posteriori and a priori information is taken. Divisions by zero avoided, the results of this paper apply to both definitions. For simplicity, we adopt the first.

[3] LoI stands for *Lattice of Information*. The reason for this nomenclature is clarified in Section 4.1.

Table 1. Operational interpretation for three traditional information-flow measures, and a new measure. The question mark indicates the value of measure.

Measure	d1: Type of question	d2: Num. questions in attack	d3: Prob. of attack successful
Shannon entropy $SE(p_S)$	*Does* $S \in \mathcal{S}'$*?*	?	S is inferred with prob. 1
Guessing entropy $NG(p_S)$	*Is* $S = s$*?*	?	S is inferred with prob. 1
Prob. of guessing $PG_n(p_S)$	*Is* $S = s$*?*	n guesses allowed	?
Prob. of guessing under $\in$ $PG_n^{\in}(p_S)$	*Does* $S \in \mathcal{S}'$*?*	n guesses allowed	?

interest. Yet, the operational interpretation of these measures also can be described in terms of questions and answers as follows.

For simplicity, assume that elements of $\mathcal{S}$ are ordered by decreasing probabilities, i.e., if $1 \leq i < j \leq |\mathcal{S}|$ then $p_S(s_i) \geq p_S(s_j)$. The *guessing entropy* of p_S is defined as $NG(p_S) = \sum_{i=1}^{|\mathcal{S}|} i \cdot p_S(s_i)$, and the *conditional guessing entropy* of p_S given a channel C_a is defined as $NG(p_S, C_a) = \sum_{o \in \mathcal{O}} p_a(o) NG(p_a(\cdot|o))$. An operational interpretation of guessing entropy is: The adversary can pose questions *Is* $S = s$*?*, for some $s \in \mathcal{S}$, to an oracle, and guessing entropy quantifies the expected number of guesses needed to learn the entire secret. Algorithmically, guessing entropy is the expected number of steps needed for the adversary to find the secret using linear search on the space of secrets.

Still assuming that the elements of $\mathcal{S}$ are in decreasing order of probabilities, the *probability of guessing* the secret in n tries is defined as $PG_n(p_S) = \sum_{i=1}^{n} p_S(s_i)$. The *conditional probability of guessing* of p_S in n tries given a channel C_a is defined as $PG_n(p_S, C_a) = \sum_{o \in \mathcal{O}} p_a(o) PG_n(p_a(\cdot|o))$. An operational interpretation of probability of guessing in n tries is: The adversary can pose questions *Is* $S = s$*?*, for some $s \in \mathcal{S}$, and the measure quantifies the probability of guessing the entire secret in n tries. Algorithmically, the probability of guessing is the chance of success by an adversary performing a linear search on the space of secrets, after n steps.

Note that the landscape of these measures is covered by varying three dimensions of their operational interpretation:
d1: the *type of question* the adversary is allowed to pose;
d2: the *number of questions (guesses)* the adversary is allowed to pose;
d3: the *probability of success*, i.e., that of the adversary inferring the secret.

Table 1 summarizes the operational interpretation of Shannon entropy, guessing entropy and probability of guessing in terms of dimensions **d1**, **d2**, and **d3**. The type of question is fixed for each measure; the other two dimensions have a dual behavior: one is fixed and the other one is quantified. In particular, Shannon entropy and guessing entropy fix the probability of guessing the secret to be 1 and quantify the number of questions necessary to do so; probability of guessing fixes the number of guesses to be n and quantifies the probability of the secret being guessed.

We add a fourth row to Table 1 for a measure whose operational interpretation is: The adversary can pose questions *Does* $S \in \mathcal{S}'$*?*, for some $\mathcal{S}' \subseteq \mathcal{S}$, to an oracle, and the measure quantifies the probability of guessing the entire secret in n tries. Algorithmically, this measure is analogous to the probability of guessing but allowing the adversary to perform a binary (rather than linear) search on the space of secrets. The *probability of guessing under* $\in$, in n tries, of a distribution p_S is defined as $PG^{\in}_n(p_S) = \max_{\mathsf{P}\in\mathtt{LoI}(\mathcal{S}),|\mathsf{P}|\leq 2^n} \sum_{\mathcal{S}'\in\mathsf{P},|\mathcal{S}'|=1} p_S(\cdot|\mathcal{S}')$. The *conditional probability of guessing under* $\in$, in n tries, of p_S given a channel C_a is defined as $PG^{\in}_n(p_S, C_a) = \sum_{o\in\mathcal{O}} p_a(o) PG^{\in}_n(p_a(\cdot|o))$.

Worth as a New Dimension. The traditional measures in Table 1 presume secrets are monolithic and equally sensitive. We relax this restriction by introducing a new dimension to the operational interpretation of measures:

d4: the *worth* the adversary extracts from a guess.

We can enrich the landscape of measures of information with new definitions that exploit the extra freedom allowed by the new dimension **d4**. As with the traditional case, for each measure we fix the type of question the adversary is allowed to pose and vary the role played by the other three dimensions. Hence we classify the measures into three groups:

- W**-measures** quantify the worth extracted from an attack when the following dimensions are fixed: (i) the number of questions that can be posed, and (ii) the required probability of success.
- N**-measures** quantify the number of guesses the adversary needs in order to succeed when the following dimensions are fixed: (i) the required probability of success, and (ii) a minimum worth-threshold to extract as measured according to a W-measure ν modeling the adversary's preferences.
- P**-measures** quantify the probability of an attack being successful when the following dimensions are fixed: (i) the number of questions that can be posed, and (ii) a minimum worth-threshold to extract as measured according to a W-measure ν modeling the adversary's preferences.

According to this classification, Shannon entropy and guessing entropy are N-measures, and probability of guessing is a P-measure (all of them implicitly using a trivial worth assignment). Table 2 organizes the measures of information worth we propose in this paper. The new table subsumes Table 1 of traditional measures.

W-measures are used to specify the fixed worth-threshold necessary to fully define P-measures and N-measures, and hence we will start our discussion with them. First we introduce a few conventions.

Assume that the set $\mathcal{S}$ of secrets follows a probability distribution p_S, and that its fields are given by set $\mathcal{F}$. Assume also that an appropriate worth assignment

Table 2. Operational interpretation for measures of information worth. The question mark indicates the value of the measure.

W-measures: quantifying worth	d1: Type of question	d2: Num. questions in attack	d3: Prob. of attack successful	d4: Worth of payoff to attacker
Worth of certainty $WCER(\omega, p_S)$	*Does* $S \in \mathcal{S}'$*?*	1 guess allowed	success with prob. 1	?
W-vulnerability $WV(\omega, p_S)$	*Does* $S \in \mathcal{S}'$*?*	1 guess allowed	? (product prob. × worth)	
Worth of exp. = $WEXP^{=}_{n,\nu}(\omega, p_S)$	*Is* $S = s$*?*	n guesses allowed	success with prob. 1	? (using W-measure ν)

N-measures: quantifying number of guesses	d1: Type of question	d2: Num. questions in attack	d3: Prob. of attack successful	d4: Worth of payoff to attacker
W-guessing entropy $WNG_{\mathsf{w},\nu}(\omega, p_S)$	*Is* $S = s$*?*	?	success with prob. 1	extracted worth w (using W-measure ν)
W-Shannon entropy $WSE_{\mathsf{w},\nu}(\omega, p_S)$	*Does* $S \in \mathcal{S}'$*?*	?	success with prob. 1	extracted worth w (using W-measure ν)

P-measures: quantifying prob. of success	d1: Type of question	d2: Num. questions in attack	d3: Prob. of attack successful	d4: Worth of payoff to attacker
W-prob. of guessing $WPG^{\in}_{\mathsf{w},n,\nu}(\omega, p_S)$	*Does* $S \in \mathcal{S}'$*?*	n guesses allowed	?	extracted worth w (using W-measure ν)

ω is provided. For an attack C_a producing observables in a set $\mathcal{O}$, the information conveyed by each $o \in \mathcal{O}$ is the information contained in the probability distribution $p_a(\cdot|o)$ that o induces on secrets. A measure of information worth is *composable* if the value of an attack can be calculated as a function of information conveyed by each observable: $\nu(\omega, p_S, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\nu(\omega, p_S(\cdot|o))$. All measures we propose in this paper are composable, but they easily extend to worst-case versions. Finally, define the *worth of a secret* $s \in \mathcal{S}$ to be the worth of learning all of its fields, i.e., $\omega(s) = \omega(\mathcal{F})$.

3.2 *W*-measures

Worth of Certainty. Consider a risk-averse adversary who is allowed to guess any part of the secret—as opposed to the secret as a whole—but who will do so only when absolutely certain the guess will succeed. To model this scenario, we note that a field is deducible with certainty from p_S if its contents is the same in every secret in the support of the distribution. Formally, the *deducible fields* from p_S are defined as $ded(p_S) = \mathcal{F} \setminus \{f \in \mathcal{F} \mid \exists s', s'' \in supp(p_S) \ : \ s'[f] \neq s''[f]\}$. For an attack C_a producing observables in a set $\mathcal{O}$, the deducible fields from each $o \in \mathcal{O}$ are those that can be inferred from the probability distribution $ded(p_a(\cdot|o))$ that o induces on secrets. The information contained in a probability distribution is defined as the worth of its deducible fields.

Definition 2 (Worth of certainty). *The* worth of certainty *of $p_{\mathcal{S}}$ is defined as $\mathit{WCER}(\omega, p_{\mathcal{S}}) = \omega(\mathit{ded}(p_{\mathcal{S}}))$. The* worth of certainty *of an attack C_a is a W-measure defined as $\mathit{WCER}(\omega, p_{\mathcal{S}}, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\, \mathit{WCER}(\omega, p_a(\cdot|o))$.*

***W*-vulnerability.** Consider an adversary who can guess a less likely structure, provided that this structure is worth enough to yield a higher overall expected gain. Formally, for every structure $\mathfrak{f} \subseteq \mathcal{F}$, we define $p_{\mathcal{S}}(\mathfrak{f})$ to be the probability that $\mathfrak{f}$ can be deduced by an adversary knowing the distribution $p_{\mathcal{S}}$: $p_{\mathcal{S}}(\mathfrak{f}) = \max_{x \in \mathcal{S}[\mathfrak{f}]} \sum_{s \in \mathcal{S}, s[\mathfrak{f}]=x} p_{\mathcal{S}}(s)$. A rational adversary maximizes the product of probability and worth, so we define W-vulnerability as follows.

Definition 3 (W-vulnerability). *The* W-vulnerability *of $p_{\mathcal{S}}$ is defined as $\mathit{WV}(\omega, p_{\mathcal{S}}) = \max_{\mathfrak{f} \subseteq \mathcal{F}} (p_{\mathcal{S}}(\mathfrak{f})\omega(\mathfrak{f}))$. The* W-vulnerability *of an attack C_a is a W-measure defined as $\mathit{WV}(\omega, p_{\mathcal{S}}, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\, \mathit{WV}(\omega, p_a(\cdot|o))$.*

Worth of expectation under =. Consider an adversary who can explore the space of secrets using brute force, i.e., by guessing the possible values of the secret, one by one. Assume that this adversary is allowed $n \geq 0$ tries. The aim is to extract as much worth as possible according to some W-measure ν modeling the adversary's preferences. This leads to the following measure.

Definition 4 (Worth of expectation under =). *Let $n \geq 0$ be the maximum number of tries allowed for the adversary. The* worth of expectation under = *of $p_{\mathcal{S}}$ is $\mathit{WEXP}^{=}_{n,\nu}(\omega, p_{\mathcal{S}}) = \max_{\mathcal{S}' \subseteq \mathcal{S}, |\mathcal{S}'| \leq n} \big(p_{\mathcal{S}}(\mathcal{S}')\omega(\mathcal{F}) + p_{\mathcal{S}}(\bar{\mathcal{S}}')\nu(\omega, p_{\mathcal{S}}(\cdot|\bar{\mathcal{S}}'))\big)$, where $\bar{\mathcal{S}}' = \mathcal{S} \backslash \mathcal{S}'$. The* worth of expectation under = *of an attack C_a is a W-measure defined as $\mathit{WEXP}^{=}_{n,\nu}(\omega, p_{\mathcal{S}}, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\, \mathit{WEXP}^{=}_{n,\nu}(\omega, p_a(\cdot|o))$.*

3.3 *N*-measures

***W*-guessing entropy.** Consider an adversary who can ask questions *Is $S = s$?* but who, instead of having to guess the secret as a whole, can fix a minimum worth $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$ to obtain according to some W-measure ν modeling the adversary's preferences. A generalized version of guessing entropy quantifies the expected number of questions to obtain a minimum worth w from such attacks.

Definition 5 (W-guessing entropy). *Let $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$ be a worth threshold quantified according to a W-measure ν. The* W-guessing entropy *of $p_{\mathcal{S}}$ is $\mathit{WNG}_{\mathsf{w},\nu}(\omega, p_{\mathcal{S}}) = \min_{\mathcal{S}' \subseteq \mathcal{S}, \nu(\omega, p_{\mathcal{S}}(\cdot|\mathcal{S}')) \geq \mathsf{w}} \big(p_{\mathcal{S}}(\bar{\mathcal{S}}') NG(p_{\mathcal{S}}(\cdot|\bar{\mathcal{S}}')) + p_{\mathcal{S}}(\mathcal{S}')(|\bar{\mathcal{S}}'| + 1)\big)$, where $\bar{\mathcal{S}}' = \mathcal{S} \backslash \mathcal{S}'$. The* W-guessing entropy *of an attack C_a is a N-measure defined as $\mathit{WNG}_{\mathsf{w},\nu}(\omega, p_{\mathcal{S}}, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\, \mathit{WNG}_{\mathsf{w},\nu}(\omega, p_a(\cdot|o))$.*

***W*-Shannon entropy.** Consider an adversary who is allowed to ask questions of the type *Does $S \in \mathcal{S}'$?* but who, instead of having to guess the entire secret, can fix a minimum worth-threshold $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$ to extract according to a W-measure ν. A generalized version of Shannon entropy quantifies the expected number of questions necessary to obtain worth w from the attacks.

Definition 6 (W-Shannon entropy). *Let* $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$ *be a worth threshold quantified according to a* W*-measure* ν*. The* W-Shannon entropy *of* p_S *is defined as* $WSE_{\mathsf{w},\nu}(\omega, p_S) = \min_{\mathsf{P} \in \mathcal{LoI}(\mathcal{S}), \forall \mathcal{S}' \in \mathsf{P}\ \nu(\omega, p_S(\cdot|\mathcal{S}')) \geq \mathsf{w}} SE(p_{\mathsf{P}})$*. The* W-Shannon entropy *of the distribution* p_S*, given an attack* C_a*, is a* N*-measure defined as* $WSE_{\mathsf{w},\nu}(\omega, p_S, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\, WSE_{\mathsf{w},\nu}(\omega, p_a(\cdot|o))$*.*

3.4 P-measures

W-Probability of Guessing. Consider an adversary allowed to pose n questions of the type *Does* $S \in \mathcal{S}'$ *?*. The following measure quantifies the chances of extracting worth $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$, as measured by some W-measure ν, from an attack. Given n questions, at most 2^n blocks can be inspected, which leads to the following mathematical definition.

Definition 7 (W-probability of Guessing). *Let* $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$ *be a worth threshold quantified according to a* W*-measure* ν*, and* $n \geq 0$ *be the maximum number of tries allowed for the adversary. The* W-probability of guessing *of* p_S *is* $WPG^{\in}_{\mathsf{w},n,\nu}(\omega, p_S) = \max_{\mathsf{P} \in \mathcal{LoI}(\mathcal{S}), |\mathsf{P}| \leq 2^n} \sum_{\mathcal{S}' \in \mathsf{P}, \nu(\omega, p_S(\cdot|\mathcal{S}')) \geq \mathsf{w}} p_S(\mathcal{S}')$*. The* W-probability of guessing *of an attack* C_a *is a* P*-measure defined as follows:* $WPG^{\in}_{\mathsf{w},n,\nu}(\omega, p_S, C_a) = \sum_{o \in \mathcal{O}} p_a(o)\, WPG^{\in}_{\mathsf{w},n,\nu}(\omega, p_a(\cdot|o))$*.*

3.5 Mathematical Properties of Measures of Information Worth

The proposed measures of information worth by definition always yield non-negative values. It is a subtler matter, however, to show that they also always yield non-negative values for leakage. Theorem 1 below shows that non-negativity of leakage holds for our measures of information worth under certain conditions. Because N-measures and P-measures have a W-measure as an input parameter to model the preferences of the adversary, we restrict consideration to W-measures presenting a consistent behavior with respect to the number of possible values for the secret. Intuitively, whenever some secret value is ruled out from the search space, the adversary's information about the secret, according to the measure, does not decrease. Formally:

Definition 8 (Monotonicity with respect to blocks). *Given a set* $\mathcal{S}$ *of secrets, a* W*-measure* ν *is said to be* monotonic with respect to blocks *if, for every worth assignment* ω*, every probability distribution* p_S *on* $\mathcal{S}$*, and all subsets (i.e., blocks)* $\mathcal{S}', \mathcal{S}''$ *of* $\mathcal{S}$ *such that* $\mathcal{S}' \subseteq \mathcal{S}''$*, it is the case that* $\nu(\omega, p_S(\cdot|\mathcal{S}')) \geq \nu(\omega, p_S(\cdot|\mathcal{S}''))$*. When* ν *quantifies uncertainty, the inequality is reversed.*

At first it might seem that monotonicity with respect to blocks would hold for every W-measure. But this is not the case. It does hold for worth of certainty, for instance, but it does not hold for W-vulnerability, as shown in the following example.

Example 1. The vulnerability of a probability distribution p_S is calculated as $V(p_S) = max_s\ p(s)$. Consider the block $S' = \{s_1, s_2, s_3, s_4\}$ of secrets, where

$p(s_1) = 1/2$ and $p(s_2) = p(s_3) = p(s_4) = 1/6$. Then $V(S') = 1/2$. Suppose that S' is split into blocks $S'' = \{s_1\}$ and $S''' = \{s_2, s_3, s_4\}$. Hence, even if $S''' \subseteq S'$, we have $V(S''') = 1/3 < V(S')$. Since traditional vulnerability is a particular case of W-vulnerability (Theorem 2), the example is also valid for the former.

In probabilistic systems, the adversary's knowledge is not tied to blocks of secrets but to probability distributions induced by observations. The concept of monotonicity is generalized accordingly.

Definition 9 (Monotonicity with respect to observations). *Given a set $\mathcal{S}$ of secrets, a measure of information worth ν is said to be* monotonic with respect to observations *if for every worth assignment ω, every probability distribution p_S on $\mathcal{S}$, and all observables $o \in \mathcal{O}$: $\nu(\omega, p_S(\cdot|o)) \geq \nu(\omega, p_S(\cdot))$. When ν quantifies uncertainty, then the inequality is reversed.*

From Example 1 it follows that W-vulnerability is not monotonic with respect to observations. It is easy to see, however, that worth of uncertainty is.

The following theorem establishes the non-negativity of leakage by showing that the adversary's information after an attack is never smaller than the a priori information.

Theorem 1. *Let $\mathcal{S}$ be a set of secrets composed by the fields in $\mathcal{F}$ and let C_a be an attack. Let ν be a W-measure that is monotonic with respect to observations, $n \geq 0$ be the number of guesses allowed for the adversary, and $0 \leq \mathsf{w} \leq \omega(\mathcal{F})$. For every distribution p_S on $\mathcal{S}$ and every worth assignment ω:*

$$\begin{aligned}
WCER(\omega, p_S, C_a) &\geq WCER(\omega, p_S) && (1)\\
WV(\omega, p_S, C_a) &\geq WV(\omega, p_S) && (2)\\
WEXP^{=}_{n,\nu}(\omega, p_S, C_a) &\geq WEXP^{=}_{n,\nu}(\omega, p_S) && (3)\\
WNG_{\mathsf{w},\nu}(\omega, p_S, C_a) &\leq WNG_{\mathsf{w},\nu}(\omega, p_S) && (4)\\
WSE_{\mathsf{w},\nu}(\omega, p_S, C_a) &\leq WSE_{\mathsf{w},\nu}(\omega, p_S) && (5)\\
WPG^{\in}_{\mathsf{w},n,\nu}(\omega, p_S, C_a) &\geq WPG^{\in}_{\mathsf{w},n,\nu}(\omega, p_S) && (6)
\end{aligned}$$

3.6 Relation with Traditional Measures

We now substantiate our claim that Shannon entropy, guessing entropy, and probability of guessing (and, in particular, vulnerability) are measures of information that ignore the worth of structures. Define the *binary worth assignment* ω_{bin} that attributes zero worth to any proper structure, i.e., $\omega_{bin}(\mathfrak{f}) = 1$ if $\mathfrak{f} = \mathcal{F}$, $\omega_{bin}(\mathfrak{f}) = 0$ if $\mathfrak{f} \subset \mathcal{F}$. Theorem 2 asserts that the traditional measures implicitly use ω_{bin} as a worth assignment, which means that only the maximal structure is deemed to be conveying relevant information. For instance, the theorem states that Shannon entropy is the particular case of W-Shannon entropy in which the adversary must perform a binary search to the maximum level of granularity, i.e., until the secret is unequivocally identified.

Theorem 2. *Let $\mathcal{S}$ be a set of secrets distributed according to p_S, and let C_a be an attack. Then the following hold:*

$$SE(p_S, C_a) = WSE_{1,WCER}(\omega_{bin}, p_S, C_a) \tag{7}$$

$$NG(p_S, C_a) = WNG_{1,WCER}(\omega_{bin}, p_S, C_a) \tag{8}$$

$$PG_n(p_S, C_a) = WEXP^{=}_{n,\nu_{null}}(\omega_{bin}, p_S, C_a) \qquad (\forall n \geq 0) \tag{9}$$

$$V(p_S|C_a) = WV(\omega_{bin}, p_S, C_a) \tag{10}$$

where ν_{null} is a W-measure such that $\nu_{null}(\omega, p_S) = 0$ for every ω and p_S.

4 Algebraic Structure for Measures of Information Worth in Deterministic Systems

4.1 Deterministic Systems and Attack Sequences

In a deterministic system $\mathcal{C}$, for each pair of high input $s \in \mathcal{S}$ and and low input $a \in \mathcal{A}$, a single output $o \in \mathcal{O}$ is produced with probability 1. Therefore each attack $a \in \mathcal{A}$ induces a partition P_a on the set of secrets, where each block $\mathcal{S}_{a,o} \in \mathsf{P}_a$ contains all secrets mapped to o when the low input to the system is a, i.e., $\mathcal{S}_{a,o} = \{s \in \mathcal{S} | \mathcal{C}(s,a) = o\}$. When the attack is clear from the context, we write $\mathcal{S}_o$ for $\mathcal{S}_{a,o}$. An attack step can be described mathematically as $\mathcal{C}(s,a) \in \mathsf{P}_a$, which is a two-phase process: (i) the adversary chooses a partition P_a on $\mathcal{S}$, corresponding to attack $a \in \mathcal{A}$, and (ii) the system responds with the block $\mathcal{S}_o \in \mathsf{P}_a$ that contains the secret.

The adversary may perform multiple attack steps for the same secret. The adversary combines information acquired in an *attack sequence* $\hat{a} = a_{t_1}, \ldots, a_{t_k}$ of k steps by intersecting the partitions corresponding to each step in the sequence, thereby obtaining a refined partition[4] $\mathsf{P}_{\hat{a}} = \bigcap_{a \in \hat{a}} \mathsf{P}_a$. Hence an attack sequence $\hat{a}$ can be modeled as a single attack where the adversary chooses the partition $\mathsf{P}_{\hat{a}}$ as the low input to the system and obtains as an observable the block to which the secret s belongs. Formally, $\mathcal{C}(s, \hat{a}) \in \mathsf{P}_{\hat{a}}$ holds.

4.2 The Lattice of Information and the Leakage from Attack Sequences

The set of all partitions on a finite set $\mathcal{S}$ forms a *complete lattice* called the *Lattice of Information* (`LoI`) [14]. The order on lattice elements is the *refinement order* $\sqsubseteq$ on partitions: $\mathsf{P} \sqsubseteq \mathsf{P}'$ iff for every $\mathcal{S}_j \in \mathsf{P}'$ there exists $\mathcal{S}_i \in \mathsf{P}$ such that $\mathcal{S}_j \subseteq \mathcal{S}_i$. The relation $\sqsubseteq$ is a partial order on the set of all partitions on $\mathcal{S}$. The *join* $\sqcup$ of two elements in the `LoI` is the intersection of partitions, and their *meet* $\sqcap$ is the transitive closure union of partitions. Given two partitions P and P', both $\mathsf{P} \sqcup \mathsf{P}'$ and $\mathsf{P} \sqcap \mathsf{P}'$ are partitions as well. We fix the deterministic system and let the elements in the `LoI` model possible executions. By controlling the

[4] The intersection of partitions is defined as $\mathsf{P} \cap \mathsf{P}' = \bigcup_{\mathcal{S}_o \in \mathsf{P}, \mathcal{S}_{o'} \in \mathsf{P}'} \mathcal{S}_o \cap \mathcal{S}_{o'}$.

low input to the system, the adversary chooses among executions, so the LoI serves as an algebraic representation of the partial order on the attack sequences the adversary can perform. Each attack sequence $\hat{a}$ corresponds to one element $\mathsf{P}_{\hat{a}}$—i.e., the partition it induces—in the LoI for $\mathcal{S}$.

An attack sequence can be seen as a path in the LoI. Each attack sequence is mapped to an element in the lattice, and by performing an attack step the adversary may obtain a finer partition on the space of secrets, therefore moving up in the lattice to a state with more information. The leakage of information from an attack sequence is, thus, the difference in the measures of information worth between the initial and final partition in the path. This definition of leakage encompasses the traditional definitions for Shannon entropy, guessing entropy, and probability of guessing.

4.3 Consistency with Respect to the LoI

The Lattice of Information has been used as an underlying algebraic structure for deterministic systems, and it provides an elegant way to reason about leakage under composition of attacks. Yasuoka and Terauchi [18] showed that orderings based on probability of guessing, guessing entropy, and Shannon entropy are all equivalent, and Malacaria [10] showed that they coincide with the refinement order in the LoI. These results establish that the traditional measures behave well with respect to the LoI: the finer a partition is, the more information (or the less uncertainty) the measures attribute to it.

All measures of information worth proposed in Section 3 behave in a similar way. That is, they are *consistent with respect to the LoI*. This is formally established in the following theorem.

Theorem 3. *Let $\mathcal{S}$ be a set of secrets composed by the fields in $\mathcal{F}$. For all* P *and* P' *in the LoI for $\mathcal{S}$, the following are equivalent:*

$$\mathsf{P} \sqsubseteq \mathsf{P}' \tag{11}$$

$$\forall \omega\ \forall p_S \quad \mathit{WCER}(\omega, p_S, \mathsf{P}) \leq \mathit{WCER}(\omega, p_S, \mathsf{P}') \tag{12}$$

$$\forall \omega\ \forall p_S\, \mathit{WV}(\omega, p_S, \mathsf{P}) \leq \mathit{WV}(\omega, p_S, \mathsf{P}') \tag{13}$$

$$\forall n\ \forall \nu\ \forall \omega\ \forall p_S \quad \mathit{WEXP}^{=}_{n,\nu}(\omega, p_S, \mathsf{P}) \leq \mathit{WEXP}^{=}_{n,\nu}(\omega, p_S, \mathsf{P}') \tag{14}$$

$$\forall \mathsf{w}\ \forall \nu\ \forall \omega\ \forall p_S \quad \mathit{WNG}_{\mathsf{w},\nu}(\omega, p_S, \mathsf{P}) \geq \mathit{WNG}_{\mathsf{w},\nu}(\omega, p_S, \mathsf{P}') \tag{15}$$

$$\forall \mathsf{w}\ \forall \nu\ \forall \omega\ \forall p_S \quad \mathit{WSE}_{\mathsf{w},\nu}(\omega, p_S, \mathsf{P}) \geq \mathit{WSE}_{\mathsf{w},\nu}(\omega, p_S, \mathsf{P}') \tag{16}$$

$$\forall \mathsf{w}\ \forall n\ \forall \nu\ \forall \omega\ \forall p_S\, \mathit{WPG}^{\in}_{\mathsf{w},n,\nu}(\omega, p_S, \mathsf{P}) \leq \mathit{WPG}^{\in}_{\mathsf{w},n,\nu}(\omega, p_S, \mathsf{P}') \tag{17}$$

where $n \geq 0$; $0 \leq \mathsf{w} \leq \omega(\mathfrak{f})$, and ν ranges over all composable W-measures that are consistent with respect to the LoI plus the worth of certainty measure WCER. In (15) *and* (16) *ν is restricted to be monotonic with respect to blocks.*

5 A Design Technique for Worth Assignments

We now outline a general technique to capture into worth assignments relevant aspects of some given scenario of interest.

The domain of worth assignments is the power set $\mathcal{P}(\mathcal{F})$ of the set $\mathcal{F}$ of fields. By endowing $\mathcal{P}(\mathcal{F})$ with the set-inclusion ordering, we obtain a (complete) *lattice of structures* $\mathtt{L}_{\mathcal{F}}$. For every structure $\mathfrak{f} \in \mathcal{P}(\mathcal{F})$ there is a partition $\mathsf{P}_{\mathfrak{f}}$, belonging to the $\mathtt{LoI}$, distinguishing structure $\mathfrak{f}$. Formally, $\mathsf{P}_{\mathfrak{f}} = \{\mathcal{S}_{s[\mathfrak{f}]=x} \mid x \in \mathcal{S}[\mathfrak{f}]\}$ where to every $x \in \mathcal{S}[\mathfrak{f}]$ corresponds the block $\mathcal{S}_{s[\mathfrak{f}]=x} = \{s \in \mathcal{S} \mid s[\mathfrak{f}] = x\}$. Proposition 1 shows that the set-inclusion ordering on structures coincides with the refinement relation on the corresponding partitions, thereby establishing that the space of structures is a sub-lattice of the $\mathtt{LoI}$.

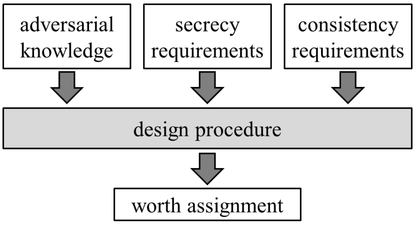

Fig. 2. Scheme of a design technique for worth assignments

Proposition 1. *For every* $\mathfrak{f}, \mathfrak{f}' \in \mathcal{P}(\mathcal{F})$*:* $\mathfrak{f} \subseteq \mathfrak{f}'$ *iff* $\mathsf{P}_{\mathfrak{f}} \sqsubseteq \mathsf{P}_{\mathfrak{f}'}$.

Hence, the space of structures $\mathtt{L}_{\mathcal{F}}$ is isomorphic to the complete lattice formed by all partitions $\mathsf{P}_{\mathfrak{f}}$ for $\mathfrak{f} \subseteq \mathcal{F}$, ordered by the refinement relation $\sqsubseteq$.

Figure 2 depicts our design technique, which constructs a worth assignment having as input the following three parameters describing a scenario of interest.

a) **Adversarial knowledge** is any relevant information the adversary knows from sources external to the system (e.g., newspapers, common-sense, other systems). As usual in QIF and privacy, adversarial knowledge is modeled as a probability distribution on the space of secrets [19–21].

b) **Secrecy requirements** reflect the protector's (i.e., the party interested in hiding the secret) interests, specifying which structures are *intrinsically sensitive* and which are only *contingently sensitive*, that is, sensitive only to the extent they possibly reveal information about other intrinsically sensitive structures. (E.g., a patient's lung cancer status may be considered intrinsically sensitive, whereas smoking habits may be considered sensitive only to the extent that they reveal information about the patient's cancer status.) Secrecy requirements are represented as a partial function from the space of structures to non-negative reals that associates every intrinsically sensitive structures with an appropriate, a priori, worth.

c) **Consistency requirements** are mathematical properties imposed on worth assignments. Non-negativity and monotonicity are considered *syntactic* consistency requirements—they depend only on the representation of secrets, not on their meaning. Syntactic requirements alone are not sufficient to guarantee the consistency of worth assignments. Often semantic requirements also need to be considered, such as the adjustments for information-theoretic predictors and computational cost from Section 2. Other examples are (i) *inclusion-exclusion consistency:* the worth of the composition of two structures is equal to the sum of their individual worths, minus the worth they share: $\omega(\mathfrak{f} \sqcup \mathfrak{f}') = \omega(\mathfrak{f}) + \omega(\mathfrak{f}') - \omega(\mathfrak{f} \sqcap \mathfrak{f}')$, and (ii) *independence:* statistically independent structures add their worth; so if $\mathsf{P}_{\mathfrak{f}'}$ and $\mathsf{P}_{\mathfrak{f}'}$ are independent then $\omega(\mathfrak{f} \sqcup \mathfrak{f}') = \omega(\mathfrak{f}) + \omega(\mathfrak{f}')$.

Once the inputs are provided, a design proceeds as follows:

1. Construct the complete lattice $\mathtt{L}_{\mathcal{F}}$ of structures.
2. Use secrecy requirements to annotate each element $\mathsf{P}_{\mathfrak{f}}$ in $\mathtt{L}_{\mathcal{F}}$, where $\mathfrak{f} \in \mathcal{P}(\mathcal{F})$ is a intrinsically sensitive structure, with the appropriate a priori worth in accordance to the protector's interests.
3. Using the adversarial knowledge, derive a probability distribution $p_{\mathcal{S}}$. Partitions in the `LoI` can be seen as random variables, so use $p_{\mathcal{S}}$ to derive the probability distribution in the elements of $\mathtt{L}_{\mathcal{F}}$.
4. Take some well established measure of information ν (e.g., guessing entropy), and for every structure $\mathfrak{f}' \in \mathcal{P}(\mathcal{F})$, update its worth according to $\omega(\mathfrak{f}') = \max_{\mathfrak{f} \in \mathcal{P}(\mathcal{F})} \nu(\mathsf{P}_{\mathfrak{f}'} | \mathsf{P}_{\mathfrak{f}})$. Repeat until all structures respect the consistency requirements.

This design technique captures the adversarial knowledge into the worth assignment, and the worth of structures will inherit the operational interpretation of the measure ν chosen in step 4. However, because the procedure depends on the probability distribution on the elements of $\mathtt{L}_{\mathcal{F}}$, certain semantic requirements only can be approximated. An example is the inclusion-exclusion principle: if it were to be preserved for all probability distributions $p_{\mathcal{S}}$, then it would be a valuation on the lattice, which is known not to exist [22].

6 Related Work

Relation with g-leakage. We start by reviewing g-leakage [13]. Given a set $\mathcal{S}$ of possible secrets and a finite, nonempty set $\mathcal{Z}$ of allowable guesses, a *gain function* is a function $g : \mathcal{Z} \times \mathcal{S} \to [0,1]$. Given a gain function g, the *prior g-vulnerability* of a probability distribution $p_{\mathcal{S}}$ is defined as $V_g(p_{\mathcal{S}}) = \max_{z \in \mathcal{Z}} \sum_{s \in \mathcal{S}} p_{\mathcal{S}}(s) g(z,s)$. Given also a channel C_a from secrets in $\mathcal{S}$ to observables in $\mathcal{O}$, the *posterior g-vulnerability* is $V_g(p_{\mathcal{S}}, C_a) = \sum_{o \in \mathcal{O}} p(o) V_g(p_a(\cdot|o))$. The g-vulnerability is converted into g-entropy by taking its logarithm: $H_g(p_{\mathcal{S}}) = -\log V_g(p_{\mathcal{S}})$ and $H_g(p_{\mathcal{S}}, C_a) = -\log V_g(p_{\mathcal{S}}, C_a)$. Finally, g-leakage is the difference between prior and posterior g-entropies: $\mathcal{L}_g(p_{\mathcal{S}}, C_a) = H_g(p_{\mathcal{S}}) - H_g(p_{\mathcal{S}}, C_a)$.

Comparing our work with g-leakage, two main points are noteworthy:

(i) g-leakage as defined in [13] cannot capture scenarios where the worth of a structure depends on the probability of that structure. Hence worth of certainty and W-Shannon entropy cannot be modeled using g-leakage.

Proposition 2. *Given a set of secrets $\mathcal{S}$ and a set of guesses $\mathcal{Z}$, there is no gain function $g : \mathcal{Z} \times \mathcal{S} \to \mathbb{R}^+$ such that, for all priors $p_{\mathcal{S}}$ on $\mathcal{S}$, and all partitions P on the `LoI` for $\mathcal{S}$, it is the case that: (i) $V_g(p_{\mathcal{S}}) = WCER(\omega, p_{\mathcal{S}})$, or (ii) $V_g(p_{\mathcal{S}}) = SE(p_{\mathcal{S}})$, or (iii) $H_g(p_{\mathcal{S}}) = SE(p_{\mathcal{S}})$.*

(ii) g-leakage and measures of information worth coincide in some scenarios, and when it happens, our approach can give practical operational interpretations to gain functions—in fact, a common criticism of the g-leakage framework concerns the challenge of identifying adequate functions for a scenario of interest. Take guessing entropy, as an example. Take an allowable guess z to be an ordered list $\Lambda(\mathcal{S}')$ of the secret elements of a subset $\mathcal{S}' \subseteq \mathcal{S}$ of secrets. A guess $\Lambda(\mathcal{S}')$ means that the adversary believes that the secret belongs to the set $\mathcal{S}'$. Moreover, in a brute-force attack the adversary would guess secrets in that same order they appear in that list. Then, for the binary worth assignment ω_{bin}[5], define a gain function $g_{\omega_{bin}}(\Lambda(\mathcal{S}'), s) = -\Lambda(\mathcal{S}')(s)$ if $s \in \mathcal{S}'$, and $g_{\omega_{bin}}(\Lambda(\mathcal{S}'), s) = -(|\bar{\mathcal{S}}'| + 1)$ otherwise. It can be shown that the W-guessing entropy captures the g-vulnerability of an adversary guided by the gain function $g_{\omega_{bin}}$, i.e., that $WNG_{1,WCER}(\omega_{bin}, p_S) = V_{g_{\omega_{bin}}}(p_S)$. However, $g_{\omega_{bin}}$ ranges over negative values, which is not allowed by the original g-vulnerability framework.[6] Fortunately we do not run into the same type of problem when using W-vulnerability, worth of expectation under $=$, and W-probability of guessing to provide operational interpretations for g-functions.

Other Related Work. Köpf and Basin [17] proposed the model for deterministic systems we extended in this paper. Shannon [23] points out the independence of the information contents with respect to its representation, and gives the first steps in trying to understand how Shannon entropy would behave in a lattice of partitions. The Lattice of Information is introduced by Landauer and Redmond [14]. Yasuoka and Terauchi [18] show the equivalence of the ordering on traditional measures, and Malacaria [10] uses the `LoI` as an algegraic foundation to unify all these orderings. Backes, Köpf and Rybalchenko [24], and Heusser and Malacaria [25] use model checkers and sat-solvers to determine the partitions induced by deterministic programs. Adão et al. [26] relax the assumption of perfect cryptography by allowing the adversary to infer a key at some (possibly computational) cost, and introduce a quantitative extension of the usual Dolev-Yao intruder model to analyze implementations of security protocols. Their work focuses on cryptography, whereas ours is applied to QIF. Askarov et al. [27] show that the possibly unbouded leakage of termination-insensitive noninterference can be mitigated by making the secret sufficiently random and large. Demange and Sands [28] point out that secrets can not always be chosen to fulfill such requirements, and they develop a framework in which "small" secrets are handled more carefully than "big" ones. They focus on preventing leakage, whereas we aim at providing rigorous information-theoretic measures for quantifying leakage.

7 Conclusion and Future Work

This paper proposed a framework to incorporate the worth of structures—possibly representing their sensitivity—into information-flow measures. We

[5] The procedure can be generalized to worth assignments other than ω_{bin}.

[6] If we try to capture W-guessing entropy using g-entropy instead of g-vulnerability, the situation becomes even worse: no gain function exists, even with negative values.

generalized Shannon entropy, guessing entropy and probability of guessing, and we proved that the generalizations are consistent with respect to the Lattice of Information for deterministic systems. We also outlined a design technique for worth assignments that captures important aspects of a scenario of interest.

We are currently refining the design technique for worth assignments to make it fully automated. We are also investigating scenarios where every attack incurs some *cost*. The resulting theory would enable the study of the trade-off between the information yielded by an attack versus cost.

Acknowledgments. The authors would like to thank Geoffrey Smith and Santosh S. Venkatesh for helpful discussions. Mário S. Alvim was a postdoctoral research associate at the Mathematics Department of the University of Pennsylvania when much of this research was performed. Support for Mário S. Alvim and Andre Scedrov comes in part from the AFOSR MURI "Science of Cyber Security: Modeling, Composition, and Measurement" as AFOSR grant FA9550-11-1-0137. Additional support for Scedrov comes from NSF grant CNS-0830949 and from ONR grant N00014-11-1-0555. Fred Schneider is supported in part by AFOSR grant F9550-06-0019, by the AFOSR MURI "Science of Cyber Security: Modeling, Composition, and Measurement" as AFOSR grant FA9550-11-1-0137, by NSF grants 0430161, 0964409, and CCF-0424422 (TRUST), by ONR grants N00014-01-1-0968 and N00014-09-1-0652, and by grants from Microsoft.

References

1. Cachin, C.: Entropy Measures and Unconditional Security in Cryptography. PhD thesis, ETH Zürich (1997) Reprint as of ETH Series in Information Security and Cryptography, vol. 1. Hartung-Gorre Verlag, Konstanz (1997) ISBN 3-89649-185-7
2. Clark, D., Hunt, S., Malacaria, P.: Quantitative information flow, relations and polymorphic types. J. of Logic and Computation 18(2), 181–199 (2005)
3. Malacaria, P.: Assessing security threats of looping constructs. In: Hofmann, M., Felleisen, M. (eds.) Proceedings of the 34th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2007, pp. 225–235. ACM (2007)
4. Malacaria, P., Chen, H.: Lagrange multipliers and maximum information leakage in different observational models. In: Proc. of the 2008 Workshop on Programming Languages and Analysis for Security (PLAS 2008), pp. 135–146. ACM (June 2008)
5. Moskowitz, I.S., Newman, R.E., Syverson, P.F.: Quasi-anonymous channels. In: Proc. of CNIS, pp. 126–131, IASTED (2003)
6. Moskowitz, I.S., Newman, R.E., Crepeau, D.P., Miller, A.R.: Covert channels and anonymizing networks. In: Jajodia, S., Samarati, P., Syverson, P.F. (eds.) Workshop on Privacy in the Electronic Society 2003, pp. 79–88. ACM (2003)
7. Chatzikokolakis, K., Palamidessi, C., Panangaden, P.: Anonymity protocols as noisy channels. Inf. and Comp. 206(2-4), 378–401 (2008)
8. Alvim, M.S., Andrés, M.E., Palamidessi, C.: Information Flow in Interactive Systems. In: Gastin, P., Laroussinie, F. (eds.) CONCUR 2010. LNCS, vol. 6269, pp. 102–116. Springer, Heidelberg (2010)
9. Massey: Guessing and entropy. In: Proceedings of the IEEE International Symposium on Information Theory, p. 204. IEEE (1994)

10. Malacaria, P.: Algebraic foundations for information theoretical, probabilistic and guessability measures of information flow. CoRR abs/1101.3453 (2011)
11. Smith, G.: On the foundations of quantitative information flow. In: de Alfaro, L. (ed.) FOSSACS 2009. LNCS, vol. 5504, pp. 288–302. Springer, Heidelberg (2009)
12. Braun, C., Chatzikokolakis, K., Palamidessi, C.: Quantitative notions of leakage for one-try attacks. In: Proceedings of the 25th Conf. on Mathematical Foundations of Programming Semantics. Electronic Notes in Theoretical Computer Science, vol. 249, pp. 75–91. Elsevier B.V. (2009)
13. Alvim, M.S., Chatzikokolakis, K., Palamidessi, C., Smith, G.: Measuring information leakage using generalized gain functions. In: Proceedings of the 25th IEEE Computer Security Foundations Symposium (CSF), pp. 265–279 (2012)
14. Landauer, J., Redmond, T.: A lattice of information. In: Proc. Computer Security Foundations Workshop VI, pp. 65–70 (June 1993)
15. Alvim, M.S., Scedrov, A., Schneider, F.B.: When not all bits are equal: Worth-based information flow. Technical report (2013), `http://ecommons.library.cornell.edu/handle/1813/33124`
16. Sweeney, L.: Uniqueness of simple demographics in the U.S. population, Carnegie Mellon University, Laboratory for International Data Privacy (2000)
17. Köpf, B., Basin, D.: Automatically deriving information-theoretic bounds for adaptive side-channel attacks. J. Comput. Secur. 19(1), 1–31 (2011)
18. Yasuoka, H., Terauchi, T.: Quantitative information flow — verification hardness and possibilities. In: Proc. 23rd IEEE Computer Security Foundations Symposium (CSF 2010), pp. 15–27 (2010)
19. Dwork, C.: Differential privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006, part II. LNCS, vol. 4052, pp. 1–12. Springer, Heidelberg (2006)
20. Ghosh, A., Roughgarden, T., Sundararajan, M.: Universally utility-maximizing privacy mechanisms. In: Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, pp. 351–360. ACM, New York (2009)
21. Alvim, M.S., Andrés, M.E., Chatzikokolakis, K., Palamidessi, C.: On the relation between differential privacy and quantitative information flow. In: Aceto, L., Henzinger, M., Sgall, J. (eds.) ICALP 2011, Part II. LNCS, vol. 6756, pp. 60–76. Springer, Heidelberg (2011)
22. Nakamura, Y.: Entropy and semivaluations on semilattices. Kodai Mathematical Seminar Reports 22(4), 443–468 (1970)
23. Shannon, C.: The lattice theory of information. IRE Professional Group on Information Theory 1(1), 105–107 (1953)
24. Backes, M., Köpf, B., Rybalchenko, A.: Automatic discovery and quantification of information leaks. In: IEEE Symposium on Security and Privacy, pp. 141–153 (2009)
25. Heusser, J., Malacaria, P.: Quantifying information leaks in software. In: Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC 2010, pp. 261–269. ACM, New York (2010)
26. Adão, P., Mateus, P., Viganò, L.: Protocol insecurity with a finite number of sessions and a cost-sensitive guessing intruder is np-complete. Theoretical Computer Science (2013) ISSN 0304-3975, `http://www.sciencedirect.com/science/article/pii/S0304397513006956`, doi:`http://dx.doi.org/10.1016/j.tcs.2013.09.015`
27. Askarov, A., Hunt, S., Sabelfeld, A., Sands, D.: Termination-insensitive noninterference leaks more than just a bit. In: Jajodia, S., Lopez, J. (eds.) ESORICS 2008. LNCS, vol. 5283, pp. 333–348. Springer, Heidelberg (2008)
28. Demange, D., Sands, D.: All secrets great and small. In: Castagna, G. (ed.) ESOP 2009. LNCS, vol. 5502, pp. 207–221. Springer, Heidelberg (2009)