Quantitative Information Flow under Generic Leakage Functions and Adaptive Adversaries

Michele Boreale, Francesca Pampaloni

Quantitative Information Flow under Generic Leakage Functions and Adaptive Adversaries*

Michele Boreale[1] and Francesca Pampaloni[2]

[1] Università di Firenze, Italy
[2] IMT - Institute for Advanced Studies, Lucca, Italy

Abstract. We put forward a model of *action-based* randomization mechanisms to analyse quantitative information flow (QIF) under generic leakage functions, and under possibly adaptive adversaries. This model subsumes many of the QIF models proposed so far. Our main contributions include the following: (1) we identify mild general conditions on the leakage function under which it is possible to derive general and significant results on adaptive QIF; (2) we contrast the efficiency of adaptive and non-adaptive strategies, showing that the latter are as efficient as the former in terms of length up to an expansion factor bounded by the number of available actions; (3) we show that the maximum information leakage over strategies, given a finite time horizon, can be expressed in terms of a Bellman equation. This can be used to compute an optimal finite strategy recursively, by resorting to standard methods like backward induction.

Keywords: quantitative information flow, adaptive attackers, information theory.

1 Introduction

Quantitative Information Flow (QIF) is a well-established approach to confidentiality analysis: the basic idea is measuring how much information flows from sensitive to public data, relying on tools from Information Theory [11,3,10,21,4,5,20,6,7].

Two major issues that arise in QIF are: what measure one should adopt to quantify the leakage of confidential data, and the relationship between adaptive and non adaptive adversaries. Concerning the first issue, a long standing debate in the QIF community concerns the relative merits of leakage functions based on Shannon entropy (see e.g. [11,5]) and min-entropy (see e.g. [21,4]); other types of entropies are sometimes considered (see e.g. [17]). As a matter of fact, analytical results for each of these types of leakage functions have been so far worked out in a non-uniform, ad hoc fashion.

Concerning the second issue, one sees that, with the notable exception of [17] which we discuss later on, QIF has so far been almost exclusively concerned with attackers that can only passively eavesdrop on the mechanism; or, at best, obtain answers in response to queries (or *actions*) submitted in a non-adaptive fashion [8]. Clearly, there

* Corresponding author: Michele Boreale, Università di Firenze, Dipartimento di Statistica, Informatica, Applicazioni (DiSIA), Viale Morgagni 65, I-50134 Firenze, Italy. E-mail: michele.boreale@unifi.it. Work partially supported by the EU project ASCENS under the FET open initiative in FP7.

are situations where this model is not adequate. To mention but two: chosen plaintext/ciphertext attacks against cryptographic hardware or software; adaptive querying of databases whose records contain both sensitive and non-sensitive fields.

In this paper, we tackle both issues outlined above. We: (a) put forward a general QIF model where the leakage function is built around a *generic uncertainty measure*; and, (b) derive several general results on the relationship between adaptive and non adaptive adversaries in this model. More in detail, we assume that, based on a secret piece of information $X \in \mathcal{X}$, the mechanism responds to a sequence of queries/actions $a_1, a_2, \ldots$ ($a_i \in Act$), adaptively submitted by an adversary, thus producing a sequence of answers/observations $Y \in \mathcal{Y}^*$. Responses to individual queries are in general probabilistic, either because of the presence of noise or by design. Moreover, the mechanism is stateless, thus answers are independent from one another. The adversary is assumed to know the distribution according to which X has been generated (the prior) and the input-output behaviour of the mechanism. An adaptive adversary can choose the next query based on past observations, according to a predefined strategy. Once a strategy and a prior have been fixed, they together induce a probability space over sequences of observations. Observing a specific sequence provides the adversary with information that modifies his belief about X, possibly reducing his uncertainty. We measure information leakage as the *average reduction in uncertainty*. We work with a generic measure of uncertainty, $U(\cdot)$. Formally, $U(\cdot)$ is just a real-valued function over the set of probability distributions on $\mathcal{X}$, which represent possible beliefs of the adversary. Just two properties are assumed of $U(\cdot)$: concavity and continuity. Note that leakage functions commonly employed in QIF, such as Shannon entropy, guessing entropy and error probability (the additive version of min-entropy) do fall in this category.

The other central theme of our study is the comparison between adaptive and the simpler non-adaptive strategies. All in all, our results indicate that, for even moderately powerful adversaries, there is no dramatic difference between the two, in terms of difficulty of analysis. A more precise account of our contributions follows.

1) We put forward a general model of adaptive QIF; we identify mild general conditions on the leakage function under which it is possible to derive general and significant results on adaptive QIF in this model.
2) We compare the difficulty of analyzing mechanisms under adaptive and non-adaptive adversaries. We first note that, for the class of mechanisms admitting a "concise" syntactic description - e.g. devices specified by a boolean circuit - the analysis problem is intractable (NP-hard), even if limited to very simple instances of the *non-adaptive* case. This essentially depends on the fact that such mechanisms can feature exponentially many actions in the syntactic size. In the general case, we show that non-adaptive finite strategies are as efficient as adaptive ones, up to an *expansion factor* in their length bounded by the number of distinct actions available. Practically, this indicates that, for mechanisms described in explicit form (e.g. by tables, like a DB) hence featuring an "affordable" number of actions available to the adversary, it may be sufficient to assess resistance of the mechanism against non-adaptive strategies. This is important, because simple analytical results are available for such strategies [8].
3) We show that the maximum leakage is the same for both adaptive and non-adaptive adversaries , and only depends on an indistinguishability equivalence relation over the set of secrets.

4) We show that maximum information leakage over a finite horizon can be expressed in terms of a Bellman equation. This equation can be used to compute optimal finite strategies recursively. As an example, we show how to do that using Markov Decision Processes (MDP's) and backward induction.

Related Work In [17], Köpf and Basin introduced an information-theoretic model of adaptive attackers for deterministic mechanisms. Their analysis is conducted essentially on the case of uniform prior distributions. Our model generalizes [17] in several respects: we consider probabilistic mechanisms, generic priors and generic uncertainty functions. More important than that, we contrast quantitatively the efficiency of adaptive and non-adaptive strategies, we characterize maximum leakage of infinite strategies, and we show how to express information leakage as a Bellman equation. The latter leads to search algorithms for optimal strategies that, when specialized to the deterministic case, are more time-efficient than the exhaustive search outlined in [17] (see Section 6).

Our previous paper [8] tackles multirun, non-adaptive adversaries, in the case of min-entropy leakage. In this simpler setting, a special case of the present framework, one manages to obtain simple analytical results, such as the exact convergence rate of the adversary's success probability as the number of observations goes to infinity.

[18,19] propose models to assess system security against classes of adversaries characterized by user-specified 'profiles'. While these models share some similarities with ours - in particular, they too employ MDP's to keep track of possible adversary strategies - their intent is quite different from ours: they are used to build and assess analysis tools, rather than to obtain analytical results. Also, the strategies they consider are tailored to worst-case adversary's utility, which, differently from our average-case measures, is not apt to express information leakage.

Alvim et al. [1] study information flow in interactive mechanisms, described as probabilistic automata where secrets and observables are seen as actions that alternate during execution. Information-theoretically, they characterize these mechanisms as channels with feedback, giving a Shannon-entropy based definition of leakage. Secret actions at each step depend on previous history, but it is not clear that this gives the adversary any ability to adaptively influence the next observation, in our sense.

Structure of the Paper Section 2 introduces the model, illustrated with a few examples in Section 3. The subsequent three sections discuss the results outlined in (2), (3) and (4) above. Section 7 contains a few concluding remarks and some directions for further research. Due to lack of space, no detailed proof is given in this version of the paper.

2 Action-based Randomization Mechanisms

2.1 Basic Definitions

Definition 1. An *action-based randomization mechanism*[1] is a 4-tuple $\mathcal{S} = (\mathcal{X}, \mathcal{Y}, Act, \{M_a : a \in Act\})$ where (all sets finite and nonempty): $\mathcal{X}, \mathcal{Y}$ and Act are re-

[1] The term *information hiding system* is sometimes found to indicate randomization mechanisms. The former term, however, is also used with a different technical meaning in the literature on watermarking; so we prefer to avoid it here.

spectively the sets of *secrets*, *observations* and *actions* (or *queries*) and for each $a \in Act$, M_a is a stochastic matrix of dimensions $|\mathcal{X}| \times |\mathcal{Y}|$.

For each action $a \in Act$, $x \in \mathcal{X}$ and $y \in Y$, the element of row x and column y of M_a is denoted by $p_a(y|x)$. Note that for each x and a, row x of M_a defines a probability distribution over $\mathcal{Y}$, denoted by $p_a(\cdot|x)$. A mechanism $\mathcal{S}$ is *deterministic* if each entry of each M_a is either 0 or 1. Note that to any deterministic mechanism there corresponds a function $f : \mathcal{X} \times Act \to \mathcal{Y}$ defined by $f(x, a) = y$, where $p_a(y|x) = 1$.

Definition 2 (Uncertainty). *Let $\mathcal{P}(\mathcal{X})$ be the set of all probability distributions on $\mathcal{X}$. A function $U : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ is an* uncertainty measure *if it is concave and continuous over $\mathcal{P}(\mathcal{X}) \subseteq \mathbb{R}^{|\mathcal{X}|}$.*

The role of concavity can be intuitively explained as follows. Suppose the secret is generated according to either a distribution p or to another distribution q, the choice depending from a coin toss, with head's probability λ. The coin toss introduces *extra randomness* in the generation process. Therefore, the overall uncertainty of the adversary about the secret, $U(\lambda \cdot p + (1-\lambda) \cdot q)$, should be *no less* than the average uncertainty of the two original generation processes considered separately, that is $\lambda U(p) + (1-\lambda)U(q)$. As a matter of fact, most uncertainty measures in QIF do satisfy concavity. Continuity is a technical requirement that comes into play only in Theorem 4.

Example 1. The following entropy functions, and variations thereof, are often considered in the quantitative security literature as measures of the difficulty or effort necessary to a passive adversary to identify a secret X, where X is a random variable over $\mathcal{X}$ distributed according to some $p(\cdot)$. All of them are easily proven to be uncertainty measures in our sense:

- *Shannon entropy*: $H(p) \stackrel{\triangle}{=} -\sum_{x \in \mathcal{X}} p(x) \log p(x)$, with $0 \log 0 = 0$ and log in base 2;
- *Error probability entropy*: $E(p) \stackrel{\triangle}{=} 1 - \max_{x \in \mathcal{X}} p(x)$;
- *Guessing entropy*: $G(p) \stackrel{\triangle}{=} \sum_{i=1}^{n-1} i \cdot p(x_i)$ with $p(x_1) \geq p(x_2) \geq \ldots \geq p(x_n)$.

A *strategy* is a partial function $\sigma : \mathcal{Y}^* \to Act$ such that $\mathrm{dom}(\sigma)$ is non-empty and prefix-closed. A strategy is *finite* if $\mathrm{dom}(\sigma)$ is finite. The *length* of a finite strategy is defined as $\max\{l \geq 0 : y^l \in \mathrm{dom}(\sigma)\} + 1$. For each $n \geq 0$ we will let $y^n, w^n, z^n, \ldots$ range over sequences in $\mathcal{Y}^n$; given $y^n = (y_1, \ldots, y_n)$ and $0 \leq j \leq n$, we will let y^j denote the first j components of y^n, $(y_1, \ldots, y_j)$. Given a strategy σ and an integer $n \geq 0$, the *truncation* of σ at level n, denoted as $\sigma\backslash n$, is the finite strategy $\sigma_{|\cup_{0 \leq i \leq n} \mathcal{Y}^i}$. A finite strategy of length l is *complete* if $\mathrm{dom}(\sigma) = \cup_{0 \leq i \leq l-1} \mathcal{Y}^i$. A strategy σ is *non-adaptive* if whenever y^n and w^n are two sequences of the same length then $\sigma(y^n) = \sigma(w^n)$ (that is, the decision of which action to play next only depends on the number of past actions); note that finite non-adaptive strategies are necessarily complete.

We note that strategies can be described as trees, with nodes labelled by actions and arc labelled by observations, in the obvious way. Any non-adaptive strategy also enjoys a simpler representation as a finite or infinite list of actions: we write $\sigma = [a_1, \ldots, a_i, \ldots]$ if $\sigma(y^{i-1}) = a_i$, for $i = 1, 2, \ldots$.

Example 2. Strategies $\sigma = [\varepsilon \mapsto a, y \mapsto b]$ and $\sigma' = [\varepsilon \mapsto a, y \mapsto b, y' \mapsto c, yy' \mapsto d]$ can be represented as in Fig. 1. Note that the tree's height is one less than the strategy's length.

2.2 Adaptive Quantitative Information Flow

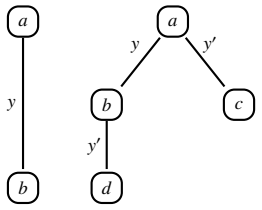

Fig. 1: Two strategy trees.

Informally, we consider an adversary who repeatedly queries a mechanism, according to a predefined *finite* strategy. At some point, the strategy will terminate, and the adversary will have collected a sequence of observations $y^n = (y_1, \ldots, y_n)$. Note that both the length n and the probability of the individual observations y_i, hence of the whole y^n, will in general depend both on X and on the strategy played by the adversary. In other words, the distribution $p(\cdot)$ of X *and* the strategy σ together induce a probability distribution on a subset of all observation sequences: the ones that may arise as a result of a complete interaction with the mechanism, according to the played strategy.

Formally, let $p(\cdot)$ be any given probability distribution over $\mathcal{X}$, which we will often refer to as the *prior*. For each finite strategy σ, we define a joint probability distribution $p_\sigma(\cdot)$ on $\mathcal{X} \times \mathcal{Y}^*$, depending on σ and on $p(\cdot)$, as follows. We let $p_\sigma(x, \varepsilon) \triangleq 0$ and, for each $j \geq 0$:

$$p_\sigma(x, y_1, \ldots, y_j, y_{j+1}) \triangleq \begin{cases} p(x) \cdot p_{a_1}(y_1|x) \cdot \cdots \cdot p_{a_j}(y_j|x) p_{a_{j+1}}(y_{j+1}|x) \\ \qquad\qquad\qquad\qquad \text{if } y^j \in \mathrm{dom}(\sigma), y^j y_{j+1} \notin \mathrm{dom}(\sigma) \\ 0 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

where in the first case $a_i = \sigma(y^{i-1})$ for $i = 1, \ldots, j+1$. In case $\sigma = [a]$, a single action strategy, we will often abbreviate $p_{[a]}(\cdot)$ as $p_a(\cdot)$. Note that the support of $p_\sigma(\cdot)$ is finite, in particular $\mathrm{supp}(p_\sigma) \subseteq \mathcal{X} \times \{y^j y : j \geq 0, y^j \in \mathrm{dom}(\sigma), y^j y \notin \mathrm{dom}(\sigma)\}$.

Let (X, Y) be a pair of random variables with outcomes in $\mathcal{X} \times \mathcal{Y}^*$, jointly distributed according to $p_\sigma(\cdot)$: here X represents the secret and Y represents the sequence of observations obtained upon termination of the strategy. We shall often use such shortened notations as: $p_\sigma(x|y^n)$ for $\Pr(X = x | Y = y^n)$, $p_\sigma(y^n)$ for $\Pr(Y = y^n)$, and so on. Explicit formulas for computing these quantities can be easily derived from the definition of $p_\sigma(\cdot)$ and using Bayes rule. We will normally keep the dependence of (X, Y) from $p(\cdot)$ and σ implicit. When different strategies are being considered at the same time and we want to stress that we are considering Y according to the distribution induced by a specific σ, we will write it as Y_σ.

Consider a prior $p(\cdot)$ and a *finite* strategy σ, and the corresponding pair of random variables (r.v.) (X, Y). We define the following quantities, expressing average uncertainty, conditional uncertainty and information gain about X, that may result from interaction according to strategy σ (by convention, we let here y^n range over sequences with $p_\sigma(y^n) > 0$):

$$\begin{aligned} U(X) &\triangleq U(p) \\ U(X|Y) &\triangleq \sum_{y^n} p_\sigma(y^n) U(p_\sigma(\cdot|y^n)) \\ I(X;Y) &\triangleq U(X) - U(X|Y). \end{aligned} \tag{1}$$

Note that, in the case of Shannon entropy, $I(X;Y)$ coincides with the familiar mutual information, traditionally measured in bits. In the case of error entropy, $I(X;Y)$ is what

is called *additive leakage* in e.g. [4] and *advantage* in the cryptographic literature, see e.g. [13] and references therein.

In the rest of the paper, unless otherwise stated, we let $U(\cdot)$ be an arbitrary uncertainty function. The following fact about $I(X;Y)$ follows from $U(\cdot)$'s concavity and Jensen's inequality, plus routine calculations on probability distributions.

Lemma 1. *$I(X;Y) \geq 0$. Moreover $I(X;Y) = 0$ if X and Y are independent.*

Given the definitions in (1), adaptive QIF can be defined quite simply.

Definition 3 (QIF **under adaptive adversaries).** *Let $\mathcal{S}$ be a mechanism and $p(\cdot)$ be a prior over $\mathcal{X}$.*

1. *For a* finite *strategy σ, let $I_\sigma(\mathcal{S}, p) \triangleq I(X;Y)$.*
2. *For an* infinite *strategy σ, let $I_\sigma(\mathcal{S}, p) \triangleq \lim_{l\to\infty} I_{\sigma\backslash l}(\mathcal{S}, p)$.*
3. *(*Maximum IF *under $p(\cdot)$) $I_\star(\mathcal{S}, p) \triangleq \sup_\sigma I_\sigma(\mathcal{S}, p)$.*

Note that $l' \geq l$ implies $I_{\sigma\backslash l'}(\mathcal{S}, p) \geq I_{\sigma\backslash l}(\mathcal{S}, p)$, hence the limit in (2) always exists. Taking the distribution that achieves the maximum leakage, we can define an analog of channel capacity.

Definition 4 (Adaptive secrecy capacity). $C(\mathcal{S}) \triangleq \sup_{p\in\mathcal{P}(\mathcal{X})} I_\star(\mathcal{S}, p)$.

2.3 Attack Trees

It is sometimes useful to work with a pictorial representation of the adversary's attack steps, under a given strategy and prior. This can take the form of a tree, where each node represents an adversary's *belief* about the secret, that is, a probability distribution over $\mathcal{X}$. The tree describes the possible evolutions of the belief, depending on the strategy and on the observations. We formally introduce such a representation below: it will be extensively used in the examples. Note that *attack* trees are different from *strategy* trees.

A *history* is a sequence $h \in (Act \times \mathcal{Y})^*$. Let $h = (a_1, y_1, \ldots, a_n, y_n)$ be such a history. Given a prior $p(\cdot)$, we define the *update of $p(\cdot)$ after h*, denoted by $p^h(\cdot)$, as the distribution on $\mathcal{X}$ defined by

$$p^h(x) \triangleq p_{\sigma_h}(x|y^n) \tag{2}$$

where $\sigma_h = [a_1, \ldots, a_n]$, provided $p_{\sigma_h}(y^n) > 0$; otherwise $p^h(\cdot)$ is undefined.

The *attack tree* induced by a strategy σ and a prior $p(\cdot)$ is a tree with nodes labelled by probability distributions over $\mathcal{X}$ and arcs labelled with pairs (y, λ) of an observation and a probability. This tree is obtained from the strategy tree of σ as follows. First, note that, in a strategy tree, each node can be identified with the unique history from the root leading to it. Given the strategy tree for σ: (a) for each $y \in \mathcal{Y}$ and each node missing an outgoing y-labelled arc, attach a new y-labelled arc leading to a new node; (b) label each node of the resulting tree by $p^h(\cdot)$, where h is the history identifying the node, if $p^h(\cdot)$ is defined, otherwise remove the node and its descendants, as well as the incoming arc; (c) label each arc from a node h to a child *hay* in the resulting tree with $\lambda = p^h_a(y)$ - to be parsed as $(p^h)_{[a]}(y)$. This is the probability of observing y under a prior $p^h(\cdot)$ when submitting action a.

The concept of attack tree is demonstrated by a few examples in the next section. Here, we just note the following easy to check facts. For each leaf h of the attack tree: (i) the leaf's label is $p^h(\cdot) = p_\sigma(\cdot|y^n)$, where y^n is the sequence of observations in h; (ii) if we let π_h be the product of the probabilities on the edges from the root to the leaf, then $\pi_h = p_\sigma(y^n)$. Moreover, (iii) each y^n s.t. $p_\sigma(y^n) > 0$ is found in the tree. As a consequence, for a *finite* strategy, taking (1) into account, the uncertainty of X given Y can be computed from the attack tree as:

$$U(X \mid Y) = \sum_{h \text{ a leaf}} \pi_h U(p^h)\,. \tag{3}$$

3 Examples

We present a few instances of the framework introduced in the previous section. We emphasize that these examples are quite simple and only serve to illustrate our main definitions. In the rest of the paper, we shall use the following notation: we let $\mathrm{u}\{x_1, \ldots, x_k\}$ denote the uniform distribution on $\{x_1, \ldots, x_k\}$.

id	ZIP	Age	Date	Desease
1	z_1	65	d_2	Hearth disease
2	z_1	65	d_2	Flu
3	z_1	67	d_2	Short breath
4	z_1	68	d_1	Obesity
5	z_1	68	d_1	Hearth disease
6	z_3	66	d_2	Hearth disease
7	z_3	67	d_2	Obesity
8	z_3	31	d_2	Short breath
9	z_2	30	d_3	Hearth disease
10	z_2	31	d_3	Obesity

Fig. 2: Medical DB of Ex. 3.

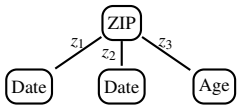

Fig. 3: Strategy tree of Ex. 3.

Example 3. An attacker gets hold of the table shown in Fig. 2, which represents a fragment of a hospital's database. Each row of the table contains: a numerical id followed by the ZIP code, age, discharge date and disease of an individual that has been recently hospitalized. The table does not contain personal identifiable information. The attacker gets to know that a certain target individual, John Doe (JD), has been recently hospitalized. However, the attacker is ignorant of the corresponding id in the table and any information about JD, apart from his name. The attacker's task is to identify JD, i.e. to find JD's id in the table, thus learning his disease. The attacker is in a position to ask a source, perhaps the hospital DB, queries concerning non sensitive information (ZIP code, age and discharge date) of any individual, including JD, and compare the answers with the table's entries.[2]

This situation can be modeled quite simply as an action-based mechanism $\mathcal{S}$, as follows. We pose: $Act = \{\mathrm{ZIP}, \mathrm{Age}, \mathrm{Date}\}$; $\mathcal{X} = \{1, \ldots, 10\}$, the set of possible id's, and $\mathcal{Y} = \mathcal{Y}_{\mathrm{ZIP}} \cup \mathcal{Y}_{\mathrm{Age}} \cup \mathcal{Y}_{\mathrm{Date}}$, where $\mathcal{Y}_{\mathrm{ZIP}} = \{z_1, z_2, z_3\}$, $\mathcal{Y}_{\mathrm{Age}} = \{30, 31, 65, 66, 67, 68\}$ and $\mathcal{Y}_{\mathrm{Date}} = \{d_1, d_2, d_3\}$. The conditional probability matrices reflect the behaviour of the source when queried about ZIP code, age and discharge date of an individual. We assume that the source is truthful, hence answers will match the entries of the table. For example, $p_{\mathrm{Age}}(y|1) = 1$ if $y = 65$ and 0 otherwise;$p_{\mathrm{ZIP}}(y|2) = 1$ if $y = z_1$, 0 otherwise; and so on. Note that this defines a *deterministic* mechanism. Finally, since the attacker has no clues about JD's id, we set the prior to be the uniform distribution on $\mathcal{X}$, $p(\cdot) = \mathrm{u}\{1, \ldots, 10\}$.

Assume now that, for some reason - maybe for the sake of privacy - the number of queries to the source about an individual is limited to two. Fig. 3 displays a possible

[2] That this is unsafe is of course well-known from database security: the present example only serves the purpose of illustration.

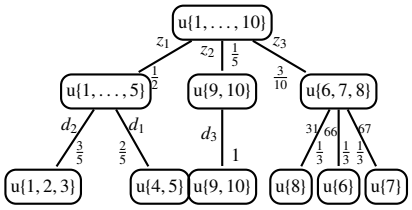

Fig. 4: The attack tree for Ex. 3.

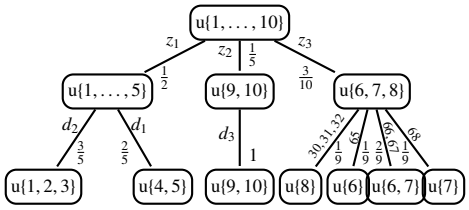

Fig. 5: The attack tree for Ex. 4. Leaves with the same label and their incoming arcs have been coalesced.

attacker's strategy σ, of length 2. Fig. 4 displays the corresponding attack tree, under the given prior. Note that the given strategy is not in any sense optimal. Assume we set $U(\cdot) = H(\cdot)$, Shannon entropy, as a measure of uncertainty. Using (3), we can compute $I_\sigma(\mathcal{S}, p) = H(X) - H(X|Y) = \log 10 - \frac{3}{10}\log 3 - \frac{2}{5} \approx 2.45$ bits. With $U(\cdot) = E(\cdot)$, the error entropy, we have $I_\sigma(\mathcal{S}, p) = E(X) - E(X|Y) = 0.5$.

Example 4 (noisy version). We consider a version of the previous mechanism where the public source queried by the attacker is not entirely truthful. In particular, for security reasons, whenever queried about age of an individual, the source adds a random offset $r \in \{-1, 0, +1\}$ to the real answer. The only difference from the previous example is that the conditional probability matrix $p_{\text{Age}}(\cdot|\cdot)$ is not deterministic anymore. For example, for $x = 1$, we have

$$p_{\text{Age}}(y|1) = \begin{cases} \frac{1}{3} & \text{if } y \in \{64, 65, 66\} \\ 0 & \text{otherwise} \end{cases}$$

(also note that we have to insert $29, 32, 64$ and 69 as possible observations into $\mathcal{Y}_{\text{Age}}$). Fig. 5 shows the attack tree induced by the strategy σ of Fig. 3 and the uniform prior in this case. If $U(\cdot) = H(\cdot)$ we obtain $I_\sigma(\mathcal{S}, p) = \log 10 - \frac{3}{10}\log 3 - \frac{8}{15} \approx 2.31$ bits; if $U(\cdot) = E(\cdot)$, instead, $I_\sigma(\mathcal{S}, p) = \frac{13}{30} \approx 0.43$.

Example 5 (cryptographic devices). We can abstractly model a cryptographic device as a function f taking pairs of a key and a message into observations, thus, $f : \mathcal{K} \times \mathcal{M} \to \mathcal{Y}$. Assume the attacker can choose the message $m \in \mathcal{M}$ fed to the device, while the key k is fixed and unknown to him. This clearly yields an action-based mechanism $\mathcal{S}$ where $\mathcal{X} = \mathcal{K}$, $Act = \mathcal{M}$ and $\mathcal{Y}$ are the observations. If we assume the observations noiseless, then the conditional probability matrices are defined by

$$p_m(y|k) = 1 \quad \text{iff} \quad f(k, m) = y\,.$$

We obtain therefore a deterministic mechanism. This is the way, for example, modular exponentiation is modeled in [17]. More realistically, the observations will be noisy, due e.g. to the presence of "algorithmic noise". For example, assume $\mathcal{Y} \subseteq \mathbb{N}$ is the set of possible Hamming weights of the ciphertexts (this is related to power analysis attacks, see e.g. [16]). Then we may set

$$p_m(y|k) = \Pr(f(k, m) + N = y)$$

where N is a random variable modelling noise. For example, in the model of DES S-Boxes considered in [8], $\mathcal{K} = \mathcal{M} = \{0, 1\}^6$, while $\mathcal{Y} = \{0, 1, 2, \ldots\}$ is the set of observations: the (noisy) Hamming weight of the outputs of the target S-Box. In this case, N is taken to be the cumulative weight of the seven S-Boxes other than the target one. It is sensible to assume this noise to be binomially distributed: $N \sim B(m, p)$, with $m = 28$ and $p = \frac{1}{2}$. See [8] for details.

4 Comparing Adaptive and Non-adaptive Strategies

Conceptually, we can classify systems into two categories, depending on the size of the set Act. Informally, the first category consists of systems with a huge - exponential, in the size of any reasonable syntactic description - number of actions. The second category consists of systems with an "affordable" number of actions. In the first category, we find, for instance, complex cryptographic hardware, possibly described via boolean circuits or other "succinct" notations (cf. the public key exponentiation algorithms considered in [17]). In the second category, we find systems explicitly described by tables, such as databases (Ex. 3 and 4) and S-Boxes (Ex.5).

4.1 Systems in Succinct Form

We argue that the analysis of such systems is in general an intractable problem, even if restricted to simple special instances of the *non-adaptive* case. We consider the problem of deciding if there is a finite strategy over a given time horizon yielding an information flow exceeding a given threshold. This decision problem is of course simpler than the problem of finding an optimal strategy over a finite time horizon: indeed, any algorithm for finding the optimal strategy can also be used to answer the first problem. We give some definitions.

Definition 5 (Systems in boolean forms). *Let t, u, v be nonnegative integers. We say a mechanism $\mathcal{S} = (\mathcal{X}, \mathcal{Y}, Act, \{M_a : a \in Act\})$ is in* (t, u, v)-boolean form *if $\mathcal{X} = \{0,1\}^t$, $Act = \{0,1\}^u$, $\mathcal{Y} = \{0,1\}^v$ and there is a boolean function $f : \{0,1\}^{t+u} \to \{0,1\}^v$ such that for each $x \in \mathcal{X}$, $y \in \mathcal{Y}$ and $a \in Act$, $p_a(y|x) = 1$ iff $f(x,a) = y$. The* size *of $\mathcal{S}$ is defined as the syntactic size of the smallest boolean formula for f.*

It is not difficult to see that the class of boolean forms coincides, up to suitable encodings, with that of deterministic systems.

Definition 6 (Adaptive Bounding Problem in succinct form, ABPS). *Given a mechanism $\mathcal{S}$ in a (t, u, v)-boolean form, a prior distribution $p(\cdot)$, $l \geq 1$ and $T \geq 0$, decide if there is a strategy σ of length $\leq l$ such that $I_\sigma(\mathcal{S}; p) > T$.*

In the following theorem, we shall assume, for simplicity, the following reasonable properties of $U(\cdot)$: if $p(\cdot)$ concentrates all the probability mass on a single element, and $q(\cdot)$ is the uniform distribution, then $0 = U(p) < U(q)$. A slight modification of the argument also works without this assumption. The theorem says that even length 1 (hence non-adaptive) strategies are difficult to assess.

Theorem 1. *Assume $U(\cdot)$ satisfies the above stated property. Then the* ABPS *is* NP*-hard, even if fixing $t = v = l = 1$, and $T = 0$.*

4.2 General Systems

The following results, which apply in general, are particularly interesting for systems with a moderate number of actions. The following theorem essentially says that, up to an expansion factor bounded by $|Act|$, non-adaptive strategies are as efficient as adaptive ones. Note that, for a strategy σ, the number of distinct actions that appear in σ is $|\text{range}(\sigma)|$.

Theorem 2. *For each finite strategy σ of length l it is possible to build a non-adaptive finite strategy σ' of length* $|\text{range}(\sigma)| \times l$*, such that* $I_{\sigma'}(\mathcal{S}, p) \geq I_\sigma(\mathcal{S}, p)$.

The intuition behind the construction of the strategy σ' is as follows. In any history induced by σ, each action can occur at most l times, and the order in which different actions appear in the history is not relevant as to the final belief that is obtained. For any history of σ to be simulated by an history of σ', it is therefore enough that the latter offers all actions offered by σ, each repeated l times. In deterministic systems, repetitions of the same action are not relevant, which leads to the following improved upper bound on the length of σ'.

Proposition 1. *If the mechanism $\mathcal{S}$ is deterministic, then the upper-bound in the previous theorem can be simplified to* $|\text{range}(\sigma)|$.

Example 6. We reconsider Ex. 3. For the adaptive strategy σ defined in Fig. 3, we have already shown that, for $U(\cdot) = H(\cdot)$, $I_\sigma(\mathcal{S}, p) \approx 2.45$. Consider now the non-adaptive strategy $\sigma' = [\text{ZIP}, \text{Date}, \text{Age}]$, which is just one action longer than σ. The corresponding attack tree is reported in Fig. 6: the final partition obtained with σ' is finer than the one obtained with σ. In fact, $I_{\sigma'}(\mathcal{S}, p) = \log 10 - \frac{2}{5} \approx 2.92 > I_\sigma(\mathcal{S}, p) \approx 2.45$.

The results discussed above are important from the point of view of the analysis. They entail that, for systems with a moderate number of actions, analyzing adaptive strategies is essentially equivalent to analyzing non-adaptive ones. The latter task can be much easier to accomplish. For example, results on asymptotic rate of convergence of non-adaptive strategies are available (e.g. [8, Th. IV.3]). They permit to analytically assess the resistance of a mechanism as the length of the considered strategies grows.

5 Maximum Leakage

In this section we show that the class of adaptive and non adaptive strategies induce the same maximum leakage. For truly probabilistic mechanisms, strategies achieving maximum leakage are in general infinite. A key notion is that of indistinguishability: an equivalence relation over $\mathcal{X}$ s.t. x and x' are indistinguishable if, no matter what strategy the adversary will play, he cannot tell them apart.

Definition 7 (Indistinguishability). *We define the following equivalence over $\mathcal{X}$:*

$$x \equiv x' \quad \textit{iff} \quad \textit{for each finite } \sigma: \; p_\sigma(\cdot|x) = p_\sigma(\cdot|x').$$

Despite being based on a universal quantification over all finite strategies, indistinguishability is in fact quite easy to characterize, also computationally. For each $a \in Act$, consider the equivalence relation defined by $x \equiv_a x'$ iff $p_a(\cdot|x) = p_a(\cdot|x')$.

Lemma 2. $x \equiv x'$ *iff for each* $a \in Act$, $p_a(\cdot|x) = p_a(\cdot|x')$. *In other words,* $\equiv$ *is* $\cap_{a \in Act} \equiv_a$.

Now, consider $\mathcal{X}/\equiv$, the set of equivalence classes of $\equiv$, and let c ranges over this set. Let $[X]$ be the r.v. whose outcome is the equivalence class of X according to $\equiv$. Note that $p(c) \stackrel{\triangle}{=} \Pr([X] = c) = \sum_{x \in c} p(x)$. We consistently extend our I-notation by defining

$$U(X \,|\, [X]) \stackrel{\triangle}{=} \sum_c p(c) U(p(\cdot \,|\, [X] = c)) \quad \text{and} \quad I(X\,;\,[X]) \stackrel{\triangle}{=} U(X) - U(X \,|\, [X]) \,.$$

More explicitly, $p(\cdot|[X] = c)$ denotes the distribution over $\mathcal{X}$ that yields $p(x)/p(c)$ for $x \in c$ and 0 elsewhere; we will often abbreviate $p(\cdot|[X] = c)$ just as $p(\cdot|c)$. Note that $I(X\,;\,[X])$ expresses the information gain about X when the attacker gets to know the indistinguishability class of the secret. As expected, this is an upper-bound to the information that can be gained playing any strategy.

Theorem 3. *$I_\star(\mathcal{S}, p) \leq I(X\,;\,[X])$.*

Proof. Fix any finite strategy σ and prior $p(\cdot)$. It is sufficient to prove that $U(X|Y) \geq U(X\,|\,[X])$. The proof exploits the concavity of U. First, we note that, for each x and y^j of nonzero probability we have (c below ranges over $\mathcal{X}/\equiv$):

$$p_\sigma(x|y^j) = \sum_c \frac{p_\sigma(x, y^j, c)}{p_\sigma(y^j)} \;=\; \sum_c p_\sigma(c|y^j)p_\sigma(x|y^j, c)\,. \tag{4}$$

By (4), concavity of $U(\cdot)$ and Jensen's inequality

$$U(p(\cdot|y^j)) \geq \sum_c p_\sigma(c|y^j)U(p_\sigma(\cdot|y^j, c))\,. \tag{5}$$

Now, we can compute as follows (as usual, y^j below runs over sequences of nonzero probability):

$$\begin{aligned} U(X|Y) &= \sum_{y^j} p_\sigma(y^j)U(p_\sigma(\cdot|y^j)) \;\geq\; \sum_{y^j,c} p_\sigma(y^j)p_\sigma(c|y^j)U(p_\sigma(\cdot|y^j, c)) && (6)\\ &= \sum_{y^j,c} p_\sigma(y^j)p_\sigma(c|y^j)U(p(\cdot|c)) \;=\; \sum_c \Big(\sum_{y^j} p_\sigma(y^j, c)\Big)U(p(\cdot|c)) && (7)\\ &= \sum_c p(c)U(p(\cdot|c)) \;=\; U(X\,|\,[X]) \end{aligned}$$

where: (6) is justified by (5); and the first equality in (7) follows from the fact that, for each x, $p_\sigma(x|y^j, c) = p(x|c)$ (once the equivalence class of the secret is known, the observation y^j provides no further information about the secret).

As to the maximal achievable information, we start our discussion from deterministic mechanism.

Proposition 2. *Let $\mathcal{S}$ be deterministic. Let $\sigma = [a_1, \ldots, a_k]$ be a non-adaptive strategy that plays all actions in Act once. Then $I_\star(\mathcal{S}, p) = I_\sigma(\mathcal{S}, p)$.*

Hence, in the deterministic case, the maximal gain in information is obtained by a trivial brute-force strategy where all actions are played in any fixed order. It is instructive to observe such a strategy at work, under the form of an attack tree. The supports of the distributions that are at the same level constitute a partition of $\mathcal{X}$: more precisely, the partition at level i ($1 \leq i \leq k$) is given by the equivalence classes of the relation $\cap_{j=1}^{i} \equiv_{a_j}$. An example of this fact is illustrated by the attack tree in Fig. 6, relative to the non-adaptive strategy [ZIP, Date, Age] for the mechanism in Ex. 3. This fact had been already observed in [17] for the restricted model considered there. Indeed, one would obtain the model of [17] by stripping the probabilities off the tree in Fig. 6.

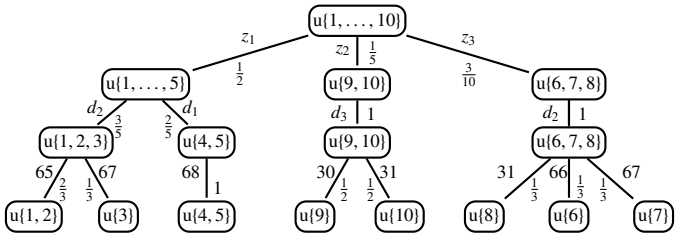

Fig. 6: The attack tree corresponding to the the non-adaptive strategy [ZIP, Date, Age] for Ex. 3.

The general probabilistic case is slightly more complicated. Essentially, any non-adaptive strategy where each action is played infinitely often achieves the maximum information gain. The next theorem considers one such strategy.

Theorem 4. *There is a total, non-adaptive strategy σ s.t. $I_\sigma(\mathcal{S}, p) = I(X; [X])$. Consequently, $I_\star(\mathcal{S}, p) = I(X\,;\,[X])$.*

Of course, as shown in the preceding section, finite adaptive strategies can be more efficient in terms of length by a factor of $|Act|$ when compared with non-adaptive ones. Concerning capacity, we do not have a general formula for the maximizing distribution. In what follows, we limit our discussion to two important cases for $U(\cdot)$, Shannon entropy and error entropy. In both cases, capacity only depends on the number K of indistinguishability classes. For guessing entropy, we conjecture that $C(\mathcal{S}) = \frac{K-1}{2}$, but at the moment a proof of this fact escapes us.

Theorem 5. *The following formulae holds, where $K = |\mathcal{X}/\equiv|$.*

- *For $U = H$ (Shannon entropy), $C(\mathcal{S}) = \log K$.*
- *For $U = E$ (Error entropy), $C(\mathcal{S}) = 1 - \frac{1}{K}$.*

Example 7. Consider the mechanism defined in Ex. 3. One has the following capacities: for $U(\cdot) = H(\cdot)$, $C(\mathcal{S}) = \log 8 = 3$, while for $U(\cdot) = E(\cdot)$, $C(\mathcal{S}) = \frac{7}{8} = 0.875$.

6 Computing Optimal Strategies

We show that, for finite strategies, $I_\sigma(\mathcal{S}, p)$ can be expressed recursively as a Bellman equation. This allows for calculation of optimal finite strategies based on standard algorithms, such as backward induction.

6.1 A Bellman Equation

Let us introduce some terminology. For each y, the *y-derivative* of σ, denoted σ_y, is the function defined thus, for each $y^j \in \mathcal{Y}^*$: $\sigma_y(y^j) \triangleq \sigma(yy^j)$. Note that if σ has length $l > 1$, then σ_y is a strategy of height $\leq l - 1$. For $l = 1$, σ_y is the empty function. Recall that according to (2), for $h = ay$, we have[3]

$$p^{ay}(x) = p_a(x|y)$$

By convention, we let $I_\sigma(\cdots)$ denote 0 when σ is empty. Moreover, we write $I_{[a]}(\cdots)$ as $I_a(\cdots)$.

[3] In terms of a given prior $p(\cdot)$ and of the matrices of $\mathcal{S}$, this can be also expressed as: $p^{ay}(x) = \frac{p_a(y|x)p(x)}{\sum_{x'} p_a(y|x')p(x')}$.

Lemma 3. *Let $p(\cdot)$ be any prior on $\mathcal{X}$. Let σ be a strategy with $\sigma(\varepsilon) = a$. Then $I_\sigma(\mathcal{S}; p) = I_a(\mathcal{S}; p) + \sum_y p_a(y) I_{\sigma_y}(\mathcal{S}; p^{ay})$.*

Let us say that a strategy σ of length l is *optimal* for $\mathcal{S}$, $p(\cdot)$ and l if it maximizes $I_\sigma(\mathcal{S}, p)$ among all strategies of length l.

Corollary 1 (Bellman-type equation for optimal strategies). *There is an optimal strategy $\sigma^\star$ of length l for $\mathcal{S}$ and $p(\cdot)$ that satisfies the following equation*

$$I_{\sigma^\star}(\mathcal{S}; p) = \max_a \{ I_a(\mathcal{S}; p) + \sum_{y:\, p_a(y)>0} p_a(y) I_{\sigma^\star_{a,y}}(\mathcal{S}; p^{ay}) \} \tag{8}$$

where $\sigma^\star_{a,y}$ is an optimal strategy of length $l-1$ for $\mathcal{S}$ and $p^{ay}(\cdot)$.

Corollary 1 allows us to employ dynamic programming or backward induction to compute optimal finite strategies. We discuss this briefly in the next subsection.

6.2 Markov Decision Processes and Backward Induction

A mechanism $\mathcal{S}$ and a prior $p(\cdot)$ induce a *Markov Decision Process* (MDP), where all possible attack trees are represented at once. Backward induction amounts to recursively computing the most efficient attack tree out of this MDP, limited to a given length. More precisely, the MDP $\mathcal{M}$ induced by $\mathcal{S}$ and a prior $p(\cdot)$ is an in general infinite tree consisting of *decision* nodes and *probabilistic* nodes. Levels of decision nodes alternate with levels of probabilistic nodes, starting from the root which is a decision node. Decision nodes are labelled with probability distributions over $\mathcal{X}$, edges outgoing decision nodes with actions, and edges outgoing probabilistic nodes with pairs (y, λ) of an observation and a real, in such a way that (again, we identify nodes with the corresponding history):

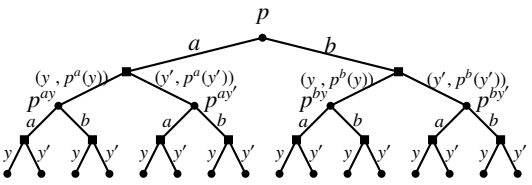

Fig. 7: The first few levels of a MDP induced by a prior $p(\cdot)$ and a mechanism with $Act = \{a, b\}$ and $\mathcal{Y} = \{y, y'\}$. Round nodes are decision nodes and squares nodes are probabilistic nodes. For the sake of space, labels of the last level of arcs and nodes are only partially shown.

- a decision node corresponding to history h is labelled with $p^h(\cdot)$, if this is defined, otherwise the node and its descendants are removed, as well as the incoming edge;
- for any pair of consecutive edges leading from a decision node h to another decision node hay, for any $a \in Act$ and $y \in \mathcal{Y}$, the edge outgoing the probabilistic node is labelled with $(y, p^h_a(y))$.

Fig. 7 shows the first few levels of such a MDP.

In order to compute an optimal strategy of length $l \geq 1$ by backward induction, one initially prunes the tree at l-th decision level (the root is at level 0) and then assigns *rewards* to all leaves of the resulting tree. Moreover, each probabilistic node is assigned an *immediate gain*. Rewards are then gradually propagated from the leaves up to the root, as follows:

– each probabilistic node is assigned as a reward the sum of its immediate gain and the *average* reward of its children, average computed using the probabilities on the outgoing arcs;
– each decision node is assigned the *maximal* reward of its children; the arc leading to the maximizing child is marked or otherwise recorded.

Eventually, the root will be assigned the maximal achievable reward. Moreover, the paths of marked arcs starting from the root will define an optimal strategy of length l. We can apply this strategy to our problem, starting with assigning rewards 0 to each leaf node h, and immediate gain $I_a(\mathcal{S}, p^h)$ to each a-child of any decision node h. The correctness of the resulting procedure is obvious in the light of Corollary 1.

In a crude implementation of the above outlined procedure, the number of decision nodes in the MDP will be bounded by $(|\mathcal{Y}| \times |Act|)^{l+1} - 1$ (probabilistic nodes can be dispensed with, at the cost of moving incoming action labels to outgoing arcs). Assuming that each distribution is stored in space $O(|\mathcal{X}|)$, the MDP can be built and stored in time and space $O(|\mathcal{X}| \times (|\mathcal{Y}| \times |Act|)^{l+1})$. This is also the running time of the backward induction outlined above, assuming $U(\cdot)$ can be computed in time $O(|\mathcal{X}|)$ (some straightforward optimizations are possible here, but we will not dwell on this). By comparison, the running time of the exhaustive procedure outlined in [17, Th.1], for deterministic systems, runs in time $O(l \times |Act|^{r^l} \times |\mathcal{X}| \times \log |\mathcal{X}|)$, where r is the maximal number of classes in any relation $\equiv_a$; since r can be as large as $|\mathcal{Y}|$, this gives a worst-case running time of $O(l \times |Act|^{|\mathcal{Y}|^l} \times |\mathcal{X}| \times \log |\mathcal{X}|)$.

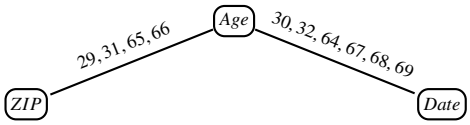

Fig. 8: A Shannon entropy optimal strategy for Ex. 4. Leaves with the same label and their incoming arcs have been coalesced.

Example 8. Applying backward induction to the mechanism of Ex. 4 with $U(\cdot) = H(\cdot)$ and $l = 2$, one gets the optimal strategy σ shown in Fig. 8, with $I_\sigma(\mathcal{S}, p) \approx 2.4$ bits.

In the general case, unfortunately, backward induction is quite memory-inefficient, even for a moderate number of observations or actions.

7 Conclusion and Further Work

We have proposed a general information-theoretic model for the analysis of confidentiality under adaptive attackers. Within this model, we have proven several results on the limits of such attackers, on the relations between adaptive and non-adaptive strategies, and on the problem of searching for optimal finite strategies.

There are several directions worth being pursued, starting from the present work. First, one would like to implement and experiment with the search algorithm described in Section 6. Adaptive querying of dataset, possibly specified via some query description language, might represent an ideal ground for evaluation of such algorithms. Second, one would like to investigate worst-case variations of the present framework: an interesting possibility is to devise an adaptive version of Differential Privacy [14,15] or one of its variants [9].

References

1. M.S. Alvim, M.E. Andrés, C. Palamidessi. Information Flow in Interactive Systems. *Journal of Computer Security*, 2011. To appear.
2. M.S. Alvim, K. Chatzikokolakis, C. Palamidessi, G. Smith. Measuring Information Leakage Using Generalized Gain Functions. *IEEE 25th Computer Security Foundations Symposium 2012*, *IEEE Computer Society*: 265-279, 2012.
3. M. Backes, B. Köpf. Formally Bounding the Side-Channel Leakage in Unknown-Message Attacks. *ESORICS 2008* , *LNCS* 5283: 517-532, 2008.
4. C. Braun, K. Chatzikokolakis, C. Palamidessi. Quantitative Notions of Leakage for One-try Attacks. *Proc. of MFPS 2009*, *Electr. Notes Theor. Comput. Sci.* 249: 75-91, 2009.
5. M. Boreale. Quantifying information leakage in process calculi. *Information and Computation* 207(6): 699-725, 2009.
6. M. Boreale, F. Pampaloni, M. Paolini. Asymptotic information leakage under one-try attacks. *Proc. of FoSSaCS 2011*, *LNCS* 6604: 396-410, 2011. Full version to appear on *Mathematical Structures in Computer Science*.
7. M. Boreale, F. Pampaloni, M. Paolini. Quantitative information flow, with a view. *Proc. of ESORICS 2011*, *LNCS* 6879: 588-606, 2011.
8. M. Boreale, F. Pampaloni. Quantitative multirun security under active adversaries. *Proc. of QEST 2012*, 158-167, 2012.
9. M. Boreale, M. Paolini. Worst- and Average-Case Privacy Breaches in Randomization Mechanisms. *IFIP TCS 2012*: 72-86, 2012.
10. K. Chatzikokolakis, C. Palamidessi, P. Panangaden. Anonimity Protocols as Noisy Channels. *Information and Computation*, 206(2-4): 378-401, 2008.
11. D. Clark, S. Hunt, P. Malacaria. Quantitative Analysis of the Leakage of Confidential Data. *Electr. Notes Theor. Comput. Sci.* 59(3), 2001.
12. T. M. Cover, J. A. Thomas. *Elements of Information Theory*, *2/e*. John Wiley & Sons, 2006.
13. Y. Dodis, A. Smith. Entropic Security and the Encryption of High Entropy Messages. *TCC 2005*, *LNCS* 3378, Springer, Heidelberg, 2005.
14. C. Dwork. Differential Privacy. *ICALP 2006. LNCS, 4052*: 1-12, 2006.
15. C. Dwork, F. McSherry, K. Nissim, A. Smith. Calibrating Noise to Sensitivity in Private Data Analysis. *Proc. of the 3rd IACR Theory of Cryptography Conference*, 2006.
16. J. Kelsey, B. Schneier, D. Wagner, C. Hall. Side Channel Cryptanalysis of Product Ciphers. *Journal of Computer Security* 8(2/3): 2000.
17. B. Köpf, D. Basin. An Information-Theoretic Model for Adaptive Side-Channel Attacks. *ACM Conference on Computer and Communications Security 2007*:286-296, 2007.
18. E. LeMay, M.D. Ford, K. Keefe, W.H. Sanders, C. Muehrcke. Model-based Security Metrics using ADversary VIew Security Evaluation (ADVISE). *Proc. of QEST 2011*, pp. 191-200, 2011.
19. M.D. Ford, P. Buchholz, W.H. Sanders. State-Based Analysis in ADVISE. *Proc. of QEST 2012*, pp. 148-157, 2012.
20. B. Köpf, G. Smith. Vulnerability Bounds and Leakage Resilience of Blinded Cryptography under Timing Attacks. *CSF 2010*: 44-56, 2010.
21. G. Smith. On the Foundations of Quantitative Information Flow. *Proc. of FoSSaCS 2009*, *LNCS* 5504: 288-302, 2009.