2016 IEEE 29th Computer Security Foundations Symposium

Axioms for Information Leakage

Mário S. Alvim* Konstantinos Chatzikokolakis† Annabelle McIver‡
Carroll Morgan§ Catuscia Palamidessi† Geoffrey Smith¶

* Computer Science Department
Universidade Federal de Minas Gerais

† CNRS, Inria, and
École Polytechnique

‡ Department of Computing
Macquarie University

§ School of Computer Science & Engineering
University of New South Wales, and Data61

¶ School of Computing & Information Sciences
Florida International University

***Abstract*—Quantitative information flow aims to assess and control the leakage of sensitive information by computer systems. A key insight in this area is that no single leakage measure is appropriate in all operational scenarios; as a result, many leakage measures have been proposed, with many different properties. To clarify this complex situation, this paper studies information leakage axiomatically, showing important dependencies among different axioms. It also establishes a completeness result about the g-leakage family, showing that any leakage measure satisfying certain intuitively-reasonable properties can be expressed as a g-leakage.**

***Index Terms*—information flow, g-vulnerability, information theory, confidentiality.**

I. Introduction

The theory of *quantitative information flow* has seen rapid development over the past decade, motivated by the need for rigorous techniques to *assess* and *control* the leakage of sensitive information by computer systems. The starting point of this theory is the modeling of a *secret* as something whose value is known to the adversary only as a *prior probability distribution* π. This immediately suggests that the "amount" of secrecy might be quantified based on π, where intuitively a uniform π would mean "more" secrecy and a biased π would mean "less" secrecy. But how, precisely, should the quantification be done?

Early work in this area (e.g., [1]) adopted classic information-theoretic measures like *Shannon-entropy* [2] and *guessing-entropy* [3]. But these can be quite misleading in a security context, because they can be arbitrarily high even if π assigns a large probability to one of the secret's possible values, giving the adversary a large chance of guessing that secret correctly in just one try. This led to the introduction of *Bayes vulnerability* [4], which is simply the maximum probability that π assigns to any of the possible values of the secret. Bayes vulnerability indeed measures a basic security threat, but it implicitly assumes an operational scenario where the adversary must guess the secret exactly, in one try. There are of course many other possible scenarios, including those where the adversary benefits by guessing a *part* or a *property* of the secret or by guessing the secret within *three tries*, or where the adversary is *penalized* for making an incorrect guess. This led to the introduction of *g-vulnerability* [5], which uses *gain functions* g to model the operational scenario, enabling specific g-vulnerabilities to be tailored to each of the above scenarios, and many others as well.[1]

This situation may however strike us as a bit of a zoo. We have a multitude of exotic vulnerability measures, but perhaps no clear sense of what a vulnerability measure ought to be. Are all the g-vulnerabilities "reasonable"? Are there "reasonable" vulnerability measures that we are missing?

The situation becomes more complex when we turn our attention to systems. We model systems as information-theoretic *channels*, and the crucial insight, reviewed in Section II-B below, is that each possible output of a channel allows the adversary to update the prior distribution π to a *posterior distribution*, where the posterior distribution itself has a probability that depends on the probability of the output. Hence a channel is a mapping from prior distributions to *distributions on posterior distributions*, called *hyper-distributions* [6].

In assessing *posterior vulnerabilities*, by which we mean the vulnerability after the adversary sees the channel output, we have a number of choices. It is natural to consider the vulnerability of each of the posterior distributions, and take the *average*, weighted by the probabilities of the posterior distributions. Or (if we are pessimistic) we might take the *maximum*. Next we can define the *leakage* caused by the channel by comparing the posterior vulnerability and prior vulnerability, either multiplicatively or additively. These choices, together with the multitude of vulnerability measures, lead us to many different leakage measures, with many different properties. Is there a systematic way to understand them? Can we bring order to the zoo?

Such questions motivate the axiomatic study that we undertake in this paper. We consider a set of axioms that characterize intuitively-reasonable properties that vulnerability measures might satisfy, separately considering axioms for prior vulnerability (Section IV) and axioms for posterior vulnerability and for the *relationship* between prior and posterior vulnerability (Section V). Addressing this relationship is an important novelty of our axiomatization, as compared with

[1] Note that *entropies* measure secrecy from the point of view of the *user* (i.e., more entropy means more secrecy), while *vulnerabilities* measure secrecy from the point of view of the *adversary* (i.e., more vulnerability means less secrecy). The two perspectives are complementary, but to avoid confusion this paper focuses almost always on the vulnerability perspective.

DOI 10.1109/CSF.2016.13