International Journal of Approximate Reasoning 69 (2016) 133–146

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Searching secrets rationally

Michele Boreale, Fabio Corradi [*]

Università di Firenze, Dipartimento di Statistica, Informatica, Applicazioni (DiSIA), Viale Morgagni 65, 50134 Firenze, Italy

ARTICLE INFO

Article history:
Received 7 May 2015
Received in revised form 24 November 2015
Accepted 27 November 2015
Available online 30 November 2015

Keywords:
Confidentiality
Quantitative information flow
Decision theory

ABSTRACT

We study quantitative information flow, from the perspective of an analyst who is interested in maximizing its expected gain in the process of learning a secret, or settling a hypothesis, represented by an unobservable X, after observing some Y related to X. In our framework, learning the secret has an associated reward, while the investigation of the set of possibilities prompted by the observation has a cost, proportional to the set's size. Approaches based on probability coverage, or on trying a fixed number of guesses, are sub-optimal in this framework. Inspired by Bayesian decision theory, we characterize the optimal behavior for the analyst and the corresponding expected gain (payoff) in a variety of situations. We argue about the importance of *advantage*, defined as the increment in expected gain after the observation if the analyst acts optimally, and representing the value of the information conveyed by Y. We characterize advantage precisely in a number of special but important instances of the framework. Applications to cryptographic systems and to familial DNA searching are examined.

1. Introduction

Broadly speaking, we refer to quantitative information flow (QIF) as the measurement of the quantity of information flowing from a unobservable random variable X to an observable Y. When expressing information as Shannon entropy [14], this quantity is just mutual information, that is, the difference between the prior and conditional entropy of X.

Computer scientists and statisticians have considered QIF from different perspectives. In the context of computer security, QIF measures expected leaks in a probabilistic system, revealing part of the secret X after some Y is observed. For a statistician, QIF corresponds to the expected reduction in uncertainty as the reward for an observation. Attackers, experimental designers and defenders are just few of the very different names assumed by the actors playing in this scene. Here, we take a somewhat neutral perspective, and simply refer to the *analyst* as someone who can expect a net gain from conditioning X on Y, in a scenario involving a cost proportional to the size of the set of possible guesses, and a reward associated with learning the secret.

In the field of quantitative security, Smith [29] has recently considered the problem of providing an adequate QIF measure for a scenario where an analyst is limited to a single guess on the candidates for the secret. An ATM withdrawing a credit card after two failed attempts at guessing the PIN illustrates the case. In this context, mutual information, which considers the global uncertainty about X before and after observing Y under a $-\log$ scale, was found to be inadequate as a measure of a QIF: in fact, the analyst's guess is now just the mode of X, so his concern is only about $V(X) = \max_x p(x)$ and $V(X|Y) = E_y[\max_x p(x|y)]$, named vulnerability and conditional vulnerability of the system, respectively. Mimicking Shannon entropy, Smith used vulnerability on the $-\log$ scale, thus obtaining an instance of Renyi's entropy called *min-entropy*.

[*] Corresponding author.
E-mail address: corradi@disia.unifi.it (F. Corradi).

http://dx.doi.org/10.1016/j.ijar.2015.11.010
0888-613X/© 2015 Elsevier Inc. All rights reserved.

In the present paper, we follow a more general approach to QIF, stemming from the tradition of Bayesian decision theory, as for example expounded in [17]. The idea is to introduce, for the problem at hand, costs associated with possible actions and a reward for learning a secret; then to derive the optimal analyst's action, that is, the one maximizing the overall expected gain. An action is just a set of possibilities that the analyst should test, or somehow further, in order to (hopefully) learn the secret, given some observable evidence. Min-entropy corresponds to the case where the reward and the costs are fixed in such a way that there is no advantage to go on testing beyond the first, most likely possibility.

In the paper, we first define a general setting from which a gain function and a QIF measure are derived (Section 2). A central role is played by *advantage*, denoted $A(X;Y)$: the difference in expected gain before and after the observation, if the analyst plays an optimal action. This represents the value, for the analyst, of the information that Y conveys about X. We then specialize the analysis by considering a fixed reward α coming from learning the secret and a unit cost for each undertaken attempt (Section 3). In this setting, we derive the optimal behavior for the analyst and characterize the resulting advantage. The behavior is shown to be more effective than both a k-tries approach with a fixed k, and the behavior based on trying guesses up to reaching a fixed probability coverage. Our results are then specialized to the important case of a non-informative (uniform) prior on the secrets, possibly in the presence of a symmetric or deterministic system (Section 4). In particular, when the reward coming from the secret equals precisely the cost of learning the secret for sure, we establish that the optimal analyst's behavior essentially corresponds to the one derived from the likelihood ratio criterion. We also consider the maximum advantage that can be obtained over all prior distributions, which is important in security contexts, that is *capacity*. We characterize capacity almost completely in the case of deterministic channels. We then examine a few applications of the proposed framework, concerning cryptographic systems and the analysis of forensic databases for familial DNA searching (Section 5). Discussion of further and related work concludes the paper (Section 6). Some detailed proofs have been confined to a separate appendix.

2. Setup

We let $\mathcal{X}$ and $\mathcal{Y}$ be finite, nonempty sets of *secrets* and *observables*, respectively. A conditional probability matrix $p_{Y|X} \in [0,1]^{\mathcal{X}\times\mathcal{Y}}$ defines the behavior of the system under observation, with $p(y|x)$ denoting the probability of the observation y when the secret is x. In the terminology of Information Theory, $p_{Y|X}$ represents the *channel* through which information about the secret flows. A prior probability p_X on $\mathcal{X}$ is assumed; we will drop the index $_X$ whenever X is clear from the context. p_X and the channel matrix $p_{Y|X}$ together give rise to a joint probability distribution on $\mathcal{X}\times\mathcal{Y}$, hence to a pair (X,Y) of input–output random variables, as expected. In many specific contexts, X and Y are not immediately related to one another, but we assume it is possible for the analyst to marginalize out all the unobserved r.v.'s in the system, apart from X. Therefore, both the prior and the conditional probability matrices are assumed to be *known to the analyst*. We will make freely use of such notational shorthand as $p(y)$ for $\Pr(Y=y)$, $p(x|y)$ for $\Pr(X=x|Y=y)$, and so on, whenever no ambiguity arises as to the underlying random variables and distributions.

Let $\mathcal{W}$ be a finite, nonempty set of *actions* the analyst can take, possibly after observing Y. Undertaking a certain action under a given state of the world/secret induces a (possibly negative) gain for the analyst, according to a given *gain function* $g : \mathcal{X}\times\mathcal{W}\to\mathbb{R}$. The *expected gain* under p_X and $w\in\mathcal{W}$ and the *maximal expected gain* under p_X are defined respectively as follows:

$$G(X;w) \stackrel{\triangle}{=} E[g(X,w)] = \sum_x g(x,w)p(x) \tag{1}$$

$$G(X) \stackrel{\triangle}{=} \max_{w\in\mathcal{W}} G(X;w)\,. \tag{2}$$

When notationally convenient, we shall use $G(X;w)$ and $G(X)$ interchangeably with $G(p;w)$ and $G(p)$, respectively, thus identifying X by its distribution p_X. In (2), a $w\in\mathcal{W}$ achieving the maximum is called a *Bayes action*. By $w^*(p)$ we indicate a Bayes action, arbitrarily chosen if there is more than one. If no ambiguity arises about p, we abbreviate $w^*(p)$ as w^*.

For $y\in\mathcal{Y}$, let $p(\cdot|y)$ denote the posterior probability distribution on $\mathcal{X}$ given $Y=y$, whenever such an event has nonzero probability, and by $G(X|y)=G(p(\cdot|y))$ the corresponding gain. The *posterior maximal expected gain*, *advantage* (under p_X) and *capacity* of the system are given by:

$$G(X|Y) \stackrel{\triangle}{=} E_y[G(X|y)] = \sum_y p(y)G(X|y) \tag{3}$$

$$A(X;Y) \stackrel{\triangle}{=} G(X|Y) - G(X) \tag{4}$$

$$C \stackrel{\triangle}{=} \sup_{p_X} A(X;Y) \tag{5}$$

where in (3) it is understood that the sum runs over y's of positive probability. General and somewhat standard results about expected gain and advantage are the following. For the sake of completeness, we report their proofs in Appendix A. We let $\mathcal{P}$ be the set of probability distributions on $\mathcal{X}$, seen as a subset of $\mathbb{R}^{|\mathcal{X}|}$.