Formal Analysis of Information Flow Using Min-Entropy and Belief Min-Entropy

Ghassen Helali, Osman Hasan, and Sofiène Tahar

Dept. of Electrical & Computer Engineering, Concordia University
1455 de Maisonneuve W., Montreal, Quebec, H3G 1M8, Canada
{helali,o_hasan,tahar}@ece.concordia.ca

Abstract. Information flow analysis plays a vital role in obtaining quantitative bounds on information leakage due to external attacks. Traditionally, information flow analysis is done using paper-and-pencil based proofs or computer simulations based on the Shannon entropy and mutual information. However, these metrics sometimes provide misleading information while dealing with some specific threat models, like when the secret is correctly guessed in one try. Min-Entropy and Belief Min-entropy metrics have been recently proposed to address these problems. But the information flow analysis using these metrics is done by simulation and paper-and-pencil approaches and thus cannot ascertain accurate results due to their inherent limitations. In order to overcome these shortcomings, we formalize Min-Entropy and Belief-Min-Entropy in higher-order logic and use them to perform information flow analysis within the sound core of the HOL theorem prover. For illustration purposes, we use our formalization to evaluate the information leakage of a cascade of channels in HOL.

Keywords: Information Flow, Min-Entropy, Belief-Min-Entropy, Information Theory, Vulnerability, Theorem Proving, Higher-order Logic, HOL4.

1 Introduction

Protecting the confidentiality of sensitive information and ensuring perfect anonymity are increasingly becoming a dire need in many fields like tele-communication, electronic payments, auctioning and voting. The information flow analysis [21] allows us to obtain quantitative estimates about information leakage, by observing the outputs and the low security inputs in a given system, and thus plays a vital role in developing secure and anonymous systems.

Various approaches for assessing the information flow have been proposed in the literature. The main idea behind the *possibilistic* approaches [1] is to use non-deterministic behaviors to model the given system. For example, the information flow analysis based on epistemic logic [8], which is a logic of knowledge and belief, and on process algebra [20], which allows us to model concurrent systems, fall under this category. The main limitation of *possibilistic* approaches

J. Iyoda and L. de Moura (Eds.): SBMF 2013, LNCS 8195, pp. 131–146, 2013.

is its failure to distinguish between systems of varying degrees of protection [6]. *Probabilistic* approaches, based on information theory and statistics, overcome this limitation and are thus considered more reliable for assessing information flow. The most commonly used probabilistic measures of information flow are Shannon's entropy [2], mutual information [3] between the sensitive input and the observable output and relative entropy [5]. It has been recently shown that using such measures sometimes leads to counter-intuitive results [22]. For example, in the case of a specific threat model where the secret is correctly guessed in one try, a random variable with high vulnerability to be guessed can have larger Shannon entropy.

In the one-try model, the adversary is given only one chance to get the value of the secret. The objective here is to maximize the probability of guessing the right value of the high input in just one try and the best strategy for her is auctionning on the element having the maximum distribution. Renyi's entropy metrics [19], i.e., Min-Entropy and Belief Min-Entropy, can deal with the above mentioned threat model more effectively and are commonly used to model and analyze the information leakage in deterministic and probabilistic systems.

Traditionally, paper-and-pencil based analysis or computer simulations have been used for quantitative analysis of information flow. Paper-and-pencil analysis does not scale well to complex systems and is prone to human error. Computer simulation, on the other hand, makes use of numerical approximations for rounding computer arithmetics, which leads to analysis inaccuracies. In order to enhance the accuracy of analysis results, formal methods have been recently proposed to be used in the safety-critical analysis domain of information flow analysis. The probabilistic model checker PRISM has been used to assist in computing the transition probabilities and capacity of the Dining cryptographers protocol [13]. However, the state-space explosion problem of model checking limits the scope of its usage in information flow analysis. For example, only the case for three cryptographers has been analyzed in [13]. These limitations can be overcome by using higher-order-logic theorem proving for the analysis of information flow. The conditional mutual information has been used to formally analyse the anonymity properties of the Dining Cryptographers protocol in the higher-order-logic theorem prover HOL4 [3]. Similarly, the information and the conditional information leakage degrees have been formalized in [17] to assess the security and anonymity protocols within the sound core of HOL4. However, to the best of our knowledge, no formalization of Min-Entropy and Belief-Min-Entropy exists in higher-order logic so far. Thus, despite their enormous applications in security-critical applications, the formal analysis of the scenarios when the secret is correctly guessed in one try is not available.

This paper presents the formalization of Min-Entropy and Belief-Min-Entropy in higher-order logic. Our formalization can be used to formally reason about the threat model where the system's vulnerability is guessed in one try by an attacker within the sound core of the HOL4 theorem prover. In this paper, we build upon the information theory foundations in HOL4 [17] mainly due to

their completeness and generic nature compared to the other formalizations of probability and information theories [4,11].

In order to illustrate the effectiveness and utilization of the proposed formalization, we utilize it to conduct the information flow analysis of channels in cascade [7]. A cascade channels topology in information theory is a commonly used linear connectivity strategy where the output of each communication node (e.g., server, router, switcher) acts as input of the next one. This structure is basically used in banking systems to ensure restorability, usability and conformity of such systems. Due to the safety-critical applications of communication systems, modeled as a cascade of channels, their accurate analysis for the worst case analysis is very important. The proposed Min-Entropy formalizations enables us to achieve this goal.

The rest of the paper is organized as follows: Section 2 provides some necessary details about the HOL theorem prover based probabilistic analysis infrastructure as well as notions of information theory that we build upon to analyze the information flow. Next, we describe the higher-order-logic definitions related to the Min-Entropy and Belief Min-Entropy theories in Section 3. We utilize these definitions in Section 4 to formally analyze the information flow. Then, we apply our new model in Section 5 to verify the Min-Entropy leakage of channels in cascade. Finally, Section 6 concludes the paper.

2 Preliminaries

This section describes the HOL4 environment as well as the formalization of probability and information theories, which we would be building upon to formalize the Min-Entropy and Belief-Min-Entropy metrics later.

2.1 HOL Theorem Prover

The HOL system is an environment for interactive theorem proving in higher order logic. Higher-order logic is a system of deduction with a precise semantics and is expressive enough to be used for the specification of almost all classical mathematics theories. In order to ensure secure theorem proving, the logic in the HOL system is represented in the strongly-typed functional programming language ML. An ML abstract data type is used to represent higher-order-logic theorems and the only way to interact with the theorem prover is by executing ML procedures that operate on values of these data types. The HOL core consists of only 5 basic axioms and 8 primitive inference rules, which are implemented as ML functions. Soundness is assured as every new theorem must be verified by applying these basic axioms and primitive inference rules or any other previously verified theorems/inference rules. The HOL system has has been used to formalize pure mathematics and verify industrial software and hardware systems.

One of the advantages of HOL is that it is not limited by the size of the state space. Large systems that cannot be verified using model checking can

still be verified by the theorem prover. Various mathematical concepts have been formalized and saved as HOL theories. Out of this useful library of HOL theories, we utilized the theories of sets, positive integers, real numbers, measure, probability and information in this paper. In fact, one of the primary motivations of selecting the HOL theorem prover for our work was to benefit from these built-in mathematical theories.

2.2 Probability Theory

Probability provides mathematical models for random phenomena and experiments. The purpose is to describe and predict relative frequencies (averages) of these experiments in terms of probabilities of events. The HOL4 utilizes the measure theory to formalize probability theory [16] and some of the foundational notions of this formalization are given below:

- **(Probability Space):** *a measure space such that the measure of the state space is* 1
- **(Independent Events):** *Two events A and B are independent iff* $p(A \cap B) = p(A)p(B)$.
- **(Random Variable):** $X : \Omega \rightarrow \mathcal{R}$ *is a random variable iff X is* $(F, \mathcal{B}(\mathcal{R}))$ *measurable where F denotes the set of events and* $\mathcal{B}$ *is the Borel sigma algebra.*
- **(Joint Probability):** *A probabilistic measure where the likelihood of two events occurring together and at the same point in time is calculated. Joint probability is the probability of event B occurring at the same time event A occurs. Its notation is* $p(A \cap B)$ *or* $p(A, B)$.
- **(Conditional Probability):** *A probabilistic measure where an event A will occur, given that one or more other events B have occurred. Its notation is* $p(A|B)$ *or* $\frac{p(A \cap B)}{p(B)}$.
- **(Expected Value):** $E[X]$ of a random variable X is its Lebesgue integral with respect to the probability measure. The following properties of the expected value have been verified in HOL4 [16]
  1. $E[X + Y] = E[X] + E[Y]$
  2. $E[aX] = aE[X]$
  3. $E[a] = a$
  4. $X \leq Y$ then $E[X] \leq E[Y]$
  5. X and Y are independent then $E[XY] = E[X]E[Y]$
- **(Variance and Covariance):** Variance and covariance have been formalized in HOL4 using the formalization of expectation. The following properties have been verified:
  1. $Var(X) = E[X^2] - E[X]^2$
  2. $Cov(X, Y) = E[XY] - E[X]E[Y]$
  3. $Var(X) \geq 0$
  4. $\forall a \in R, Var(aX) = a^2 Var(X)$
  5. $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

The above mentioned definitions and properties have been utilized to formalize the foundations of information theory in HOL4 [16].

2.3 Information Theory

Information theory [14,5] is used in many fields of engineering and computer science, such as signal processing, data compression, storing and communicating data to quantify information. Recently, it found an enormous application in the domains of cryptography and information flow analysis [23]. Various information theoretic notions, such as the *entropy*, *the mutual information*, *the relative entropy*, *the conditional entropy* and *the Renyi's entropy*, are used to reason about the uncertainty of a random variable.

Let X and Y denote discrete random variables, with x and y and $\mathcal{X}$ and $\mathcal{Y}$ denoting their specific values and set of all possible values, respectively. Similarly, the probability of X and Y being equal to x and y is denoted by $p(x)$ and $p(y)$, respectively, their joint probability is denoted by $p(x, y)$. Now, the widely used information theoretic measures can be defined as:

- **(The Shannon Entropy):** It measures the uncertainty of a random variable

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) log\, p(x)$$

- **(The Conditional Entropy):** It measures the amount of uncertainty of X when Y is known

$$H(X|Y) = -\sum_{y \in \mathcal{Y}} p(y) \sum_{x \in \mathcal{X}} p(x|y) log\, p(x|y)$$

- **(The Mutual Information):** It represents the amount of information that has been leaked

$$I(X;Y) = I(Y;X) = H(X) - H(X|Y)$$

- **(The Relative Entropy or Kullback Leiber Distance):** It measures the inaccuracy or information divergence of assuming that the distribution is q when the true distribution is p

$$D(p\|q) = \sum_{x \in \mathcal{X}} p(x) log\, \frac{p(x)}{q(x)}$$

- **(The Guessing Entropy):** It measures the expected number of tries required to guess the value of X optimally

$$G(X) = \sum_{1 \leq i \leq n} i p(x_i)$$

- **(The Rnyi Entropy):** It is related to the difficulty of guessing the value of X

$$H_\alpha(X) = \tfrac{1}{1-\alpha}\, log\, (\sum_{x \in \mathcal{X}} P[X = x]^\alpha)$$

The above measures are used to analyze the information flow from different aspects. Entropy, Mutual Information and Relative Entropy, operate over the quantity of information and the degree of uncertainty while the Guessing Entropy determines the number of attempts to decrypt a secret. Mhamdi [15] and Coble [3] formalized the notions of Entropy, Conditional Entropy, Relative Entropy and Mutual Information in HOL4, while Hölzl [11] formalized the same concepts in Isabelle/HOL.

3 Formalization of Min-Entropy and Belief Min-Entropy

Information theoretic measures of Min-Entropy and Belief Min-Entropy overcome the limitations of Shannon's entropy in evaluating the security of guessing the secret in *one try* [23]. We explain these measures along with their corresponding higher order-logic formalizations in this section. In the following subsections, X, Y and B denote the random variables that model the high input (the secret), the output (the observable) and the attacker's belief about the system behavior (the extra knowledge), respectively, and p and q denote probability spaces.

3.1 Formalization of Min-Entropy

The Min Entropy H_∞ of a random variable X is a special case from the Rényi Entropy when $\alpha = \infty$.

Definition 1 *(The Min-Entropy).*
The Min-Entropy of a random variable X is given by

$$H_\infty(X) = -log \max_{x \in \mathcal{X}} p(x)$$

This can be formalized in HOL4 as follows:

```
⊢ ∀ X p.
  min_entropy X p =
  - log (extreal_max_set (IMAGE
  (λx. distribution p X {x}) (IMAGE X (p_space p))))
```

In this definition, the function `extreal_max_set` returns the maximum of a given set, `IMAGE f s` returns the image of a given set s by a function f and `p_space p` is the state space of the Ω of the probability space p.

It can be observed from the above definition that the Min-Entropy measure is primarily the negative logarithm of the vulnerability, or in other words, the worst-case probability that an adversary A can guess the secret correctly in one try:

$$H_\infty(X) = -log(V(X)) = -log(\max_{x \in \mathcal{X}} P[X = x]).$$

The Min-Entropy measures the initial uncertainty only and the remaining uncertainty can be quantified by the conditional Min-Entropy.

Definition 2 *(The Conditional Min-Entropy).*
Observing the output Y, the probability of guessing the secret X is

$$H_\infty(X|Y) = -log(\sum_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} P[Y=y]P[X=x|Y=y])$$

This can be formalized in HOL4 as follows:

```
⊢ ∀ X Y p.
  conditional_min_entropy p X Y =
  - log ∑ (λy. extreal_max_set
      (IMAGE (λx. distribution p Y{y}*
         conditional_distribution p X Y ({x},{y})) (X(Ω))))
```

In the above definition, we utilized the `conditional_distribution p X Y` that refers to $P(X|Y)$. This quantity relates two behaviors, i.e., the input X and the output Y, and this makes the Conditional Min-Entropy a good measure to map the remaining uncertainty, which is nothing but the probability of guessing the secret input having the observable.

3.2 Formalization of Belief Min-Entropy

The Belief Min-Entropy allows us to deal with the attacker's extra knowledge or beliefs about the system behavior. This measure is actually a refinement of the Min-Entropy since it takes into account another parameter, i.e., *belief*, that is expected to increase the reliability of the analysis.

Let p_ρ and p_β denote the distributions related to the system behavior and the adversary's belief, respectively. Given an additional information $B = b$, the adversary chooses a value having the maximal conditional probability according to her belief, that is a value $x' \in \Gamma_b$, such that $\Gamma_b = argmax_{x\in\mathcal{X}}\ p_\beta(x|b)$, and $argmax_{x\in\mathcal{X}}\ p_\beta(a|b)$ returns the elements from $\mathcal{A}$ having the maximal conditional-distribution. In case of more than one value of A with the maximal conditional probability, the attacker uniformly and randomly picks a single element from Γ_b.

Definition 3 *(The Belief Min-Entropy).*
Let X be the input random variable and B the adversarys extra knowledge about X. Then the Belief Min-Entropy of X, denoted $H_\infty(X:B)$, is defined as

$$H_\infty(X:B) = -log(\sum_{b\in\mathcal{B}} \frac{1}{|\Gamma_b|}\, p(b) \sum_{x\in\mathcal{X}} p(x|b))$$

In order to formalize the Belief Min-Entropy in HOL4, we first define the belief vulnerability, which can be extended to obtain the Belief Min-Entropy by applying the converse logarithm.

```
⊢ ∀ p1 p2 X B.
  belief_vulnerability p1 p2 X B =
  SIGMA (λb. 1/|belief_set p1 p2 X B b| *
     (distribution p1 B {b}) *
        (SIGMA (λx. conditional_distribution p1 X B ({x},{b}))
        (belief_set p1 p2 X B b)))
        (B(Ω1))
```

where the function `belief_set p1 p2 X B b` models Γ_b in HOL4 and $\Omega 1$ refers to p_space p1. Now, in order to model the Belief Min-Entropy, we need to define the relationship between the attacker's belief and the observable output. The belief b is compatible with the observation y, if there exists an input $x \in \Gamma_b$ verifying $p_\rho(y|x) > 0$ and in this case, the attacker is able to choose the appropriate values for guessing the secret.$\Gamma_{b,y}$ denotes the set of the possibilities that the adversary can choose and is defined as follows:

$$\Gamma_{b,y} = \begin{cases} arg\max\limits_{x \in \mathcal{X}} \; p_\beta(x|b,y) & \text{if } b \text{ and } y \text{ are compatible} \\ arg\max\limits_{x \in \mathcal{X}} \; p_\beta(x|y) & \text{otherwise} \end{cases} \tag{1}$$

The above definition is formalized as the HOL4 function `belief_conditionned_set`, which we will use later to model the remaining uncertainty that will be a function of the conditional belief vulnerability.

Definition 4 *(The Conditional Belief Vulnerability).*

$$V(X|Y:B) = \sum_{y \in \mathcal{Y}} \sum_{b \in \mathcal{B}} p_\rho(y,b) \frac{1}{|\Gamma_{b,y}|} \sum_{x \in \Gamma_b \, p(x|y,b)}$$

The above definition can be formalized in HOL4 as follows:

```
⊢ ∀ p1 p2 X B Y. conditional_belief_vulnerability p1 p2 X B Y =
   ∑y ∑b joint_distribution p1 B Y ({b},{y}) * 1/|Γb,y| *
   ∑x∈Γb,y belief_conditional_distribution p1 X Y B ({x},{y},{b})
```

Now, we can apply the converse logarithm to get the conditional Belief Min-Entropy.

$$H_\infty(X|Y:B) = -log(V(X|Y:B))$$

Based on the previous measures, we define the information leakage that determines how much information has been leaked from the input to the output.

information leakage = initial uncertainty - remaining uncertainty

Next, we will use the definitions, presented in this section, to formally reason about their classical properties, which in turn allow us to conduct formal information flow analysis with the HOL4 theorem prover.

4 Formal Analysis of Information Flow

The main focus of this paper is on the analysis of the threat model of guessing the critical information in one try, which is usually considered as the worst case scenario and cannot be handled by the Shannon entropy as we mentioned earlier. In this section, we formally verify that the definitions, presented in the previous section, can handle this particular model.

In regards to information flow analysis, Min-Entropy allows us to measure uncertainties. The following theorem provides a lower bound to the initial uncertainty.

Theorem 1 *(Lower Bound of the Min-Entropy).*

$$\begin{aligned}&\vdash \forall\ \texttt{X p b. FINITE } (\Omega) \wedge \Omega \neq \emptyset \wedge \texttt{random\_variable X p Borel} \wedge \\ &\quad (\forall \texttt{x. x} \in \texttt{X}(\Omega) \Rightarrow \texttt{(distribution p X \{x\})} \leq \tfrac{1}{2^{\texttt{b}}})) \wedge \\ &\quad (\forall \texttt{x. x} \in \Omega \Rightarrow \{\texttt{x}\} \in \texttt{events p}) \wedge \\ &\quad X(\Omega) \in \texttt{subsets Borel}) \Rightarrow \\ &\quad \texttt{b} \leq \texttt{(min\_entropy X p)}\end{aligned}$$

where Ω = `p_space p`. If this initial uncertainty is uniformly distributed over the input set $\mathcal{X}$, then the initial uncertainty is equal to $|\mathcal{X}|$:

Theorem 2 *(Initial Uncertainty for Uniform Distribution).*

$$\begin{aligned}&\vdash \forall \texttt{p X. FINITE } (\Omega) \wedge \\ &\quad \texttt{random\_variable X p Borel} \wedge \\ &\quad \forall \texttt{x. x} \in \texttt{X}(\Omega) \Rightarrow \texttt{distribution p X \{x\} = 1 / |X}(\Omega)\texttt{|} \\ &\quad \Rightarrow \texttt{min\_entropy X p = log |X}(\Omega)\texttt{|}\end{aligned}$$

The first assumption, in the above theorems, is required because the maximum of a set is well-defined for finite sets only.

Another useful aspect related to information leakage is the remaining uncertainty that represents the model of the aposteriori behavior. If a program is deterministic and the initial distribution is uniformly distributed, then its information leakage depends on the output set only. This result can be formally verified as the following theorem:

Theorem 3 *(Information Leakage of Deterministic Program).*

$$\begin{aligned}&\vdash \forall\ \texttt{X Y p c. (}\forall \texttt{x. x} \in \texttt{X}(\Omega) \Rightarrow \texttt{distribution p X \{x\} = } \tfrac{1}{|X(\Omega)|}\texttt{)} \wedge \\ &\quad \texttt{deterministic\_cond Y c} \Rightarrow \\ &\quad \texttt{information\_leakage p X Y = log (|Y}(\Omega)\texttt{|)}\end{aligned}$$

where the assumptions model the determinism condition and the uniform distribution. Next, we analyze the information flow considering the attacker's belief. For this purpose, we include another random variable B that models the adversary's extra knowledge about the high input. Under the condition of a total inaccurate belief, the following theorem holds:

Theorem 4 *(Initial Uncertainty of Total Inaccurate Belief).*

```
⊢ ∀A B sp ev p1 p2. FINITE (p_space (sp,ev,p2)) ∧
   FINITE (p_space (sp,ev,p1)) ∧
   ∀a b. (a,b) ∈ totally_inaccurate_belief_set sp ev p1 p2 A B ⇒
   belief_min_entropy sp ev p1 p2 A B = +∞
```

According to the above theorem, when the attacker has no information about the secret input, the initial vulnerability of the system tends to zero. The proof of this result is based on the Bayes' rule and our definition of the Belief Min-Entropy.

The following theorem verifies that the conditional Min-Entropy is always less then or equal to the Belief Min-Entropy:

Theorem 5 *(Min-Entropy and Belief Min-Entropy).*

```
⊢ ∀X B sp ev p1 p2.
   ∀x b. (b ∈ B(Ω1)) ∧
      (belief_set (sp,ev,p1) (sp,ev,p2) X B b ≠ ∅) ∧
      (x ∈ (belief_set (sp,ev,p1) (sp,ev,p2) X B b)) ∧
   conditional_distribution (sp,ev,p1) B X ({b},{x}) ≤ 1/|B(Ω1)| ⇒
   min_entropy A (sp,ev,p1) ≤ belief_min_entropy sp ev p1 p2 X B
```

The interpretation of the previous result is that the vulnerability of a system is greater in the presence of the extra knowledge. Similarly, the following theorem provides the belief initial uncertainty in the deterministic case.

Theorem 6 *(Deterministic Belief Initial Uncertainty).*

```
⊢ ∀X B sp ev p1 p2 c.
   ∀x b. x ∈ belief_set (sp,ev,p1) (sp,ev,p2) X B b ∧
      b ∈ B(Ω1) ∧
   ∀x. (x ∈ belief_set (sp,ev,p1) (sp,ev,p2) X B b) ⇒
      distribution (sp,ev,p1) X {x} = 1/|X(Ω1)| ∧
   events (sp,ev,p1) = POW (Ω1) ∧
   deterministic_cond B c ⇒
   log |X(Ω1)|/|B(Ω1)| ≤ belief_min_entropy sp ev p1 p2 A B
```

Next, just like in the case of Min-Entropy, we verify that the remaining belief uncertainty is lower bounded by conditional Min-Entropy joint to the adversary's belief, i.e. $H_\infty(A|O,B) \leq H_\infty(A|O:B)$, which can be expressed as the following HOL4 theorem:

Theorem 7 *(Lower Bound for Belief Remaining Uncertainty).*

```
⊢ ∀X B Y p1 p2. FINITE (Ω) ∧ random_variable X p1 Borel ∧
   random_variable B p1 Borel ∧ random_variable Y p1 Borel ∧
   ∀x. x ∈ (Ω) ⇒ {x} ∈ events p1
   ⇒ conditional_joint_min_entropy p1 X B Y ≤
      conditional_belief_min_entropy p1 p2 X B Y
```

Thus, the belief remaining uncertainty under the deterministic conditions is bounded by $log(\frac{|\mathcal{A}|}{|\mathcal{O}|.|\mathcal{B}|})$. Now we can formally verify the following result in HOL4.

Theorem 8 *(Deterministic Remaining Belief Uncertainty).*

$\vdash \forall$`X Y B p q c c'. FINITE` $\Omega \wedge \Omega \neq \emptyset \wedge$
$\forall$`x b y. x` $\in$ `belief_conditionned_set p q X B Y b y` $\wedge$
$\forall$`b. b` $\in$ `B(`Ω`)` $\wedge \forall$`y. y` $\in$ `Y(`Ω`)` $\wedge$
$\forall$`x. x` $\in \Omega \Rightarrow$ `{x}` $\in$ `events p` $\wedge$
$\forall$`x. x` $\in$ `x(`Ω`)` $\Rightarrow$ `distribution p X {x} =` $\frac{1}{|X(\Omega)|} \wedge$
`deterministic_cond Y c` $\wedge$ `deterministic_cond B c'` $\Rightarrow$
`log` $(\frac{|X(\Omega)|}{|Y(\Omega)|.|B(\Omega)|}) \leq$ `conditional_belief_min_entropy p q X B Y`

where `belief_conditionned_set p q X B Y b y` $= \Gamma_{b,y}$ denotes the set of possible adversarys choices according to her belief and low observation.

The proof of the above theorem is primarily based on the Min-Entropy properties under deterministic conditions. Finally, Theorems 6 and 8 can be used to reason about the belief information leakage for deterministic programs.

$$log|\mathcal{Y}| \leq IL_{\infty}(X;(Y:B))$$

From the above result, we conclude that the belief behavior helps the adversary in choosing more reliable initial knowledge based on the observations. The above mentioned properties have been verified before [9] but the main novelty of our work was to re-verify these results using an interactive theorem prover. Based on the soundness of theorem proving, the formally verified theorems are guaranteed to be accurate and contain all the required assumptions. Moreover, these formally verified results can be built upon to reason about information flow analysis of various applications within the sound core of a theorem prover. For illustration purposes, the information leakage of cascade of channels is formally analyzed in the next section. These added advantages have been attained at the cost of human effort in formalizing and interactively verifying the above mentioned results. The proof script [10] is composed of 3400 lines of code and took about 1000 man-hours of development time.

5 Application: Channels in Cascade

A channel [7] is a triplet $(\mathcal{A}, \mathcal{B}, \mathcal{C}_{\mathcal{AB}})$, where $\mathcal{A}$ is a finite set of the critical inputs, $\mathcal{B}$ is the observable output and $\mathcal{C}_{\mathcal{AB}}$ is the channel matrix representing the transitional probabilities from the input to the output of the channel. The channels are frequently connected in a cascade manner such that the outputs of the previous stage act as the input to the next one. In cascaded channels, the final output is produced in n steps, where n represents the number of cascaded channels.

The major goal of this section is to formally reason about the information flow of channels in cascade and analyze the information leakage in such systems.

We will first formalize the notions of channels and cascade of channels in higher-order logic. These definitions, along with our formally verified results of the previous section, will then be used to formally reason about the measure of quantity of information and the information leakage of a two cascaded channel model.

5.1 Formalization of Channels and Cascade of Channels

A channel can be formalized in HOL4 using the following function:

Definition 5 *(Channel).*

```
⊢ ∀X Y p f. channel p X Y f =
    random_variable X p Borel ∧
    random_variable Y p Borel ∧
    ∀x y. x ∈ X(Ω) ∧
    y ∈ Y(Ω) ∧
    f(x,y) = conditional_distribution p Y X ({y},{x})
```

The predicate `channel` accepts a probability space `p`, the random variables `X` and `Y` representing the finite sets of the critical inputs and the observable outputs, respectively, and a function `f` that models the channel matrix $\mathcal{C}_{\mathcal{AB}}$ in terms of the conditional probabilities of obtaining the output b such that the input is a.

Now the behavior of a cascade of two channels, i.e., $(\mathcal{X}, \mathcal{Z}, \mathcal{C}_{\mathcal{XZ}})$ and $(\mathcal{Z}, \mathcal{Y}, \mathcal{C}_{\mathcal{ZY}})$, is equivalent to the channel $(\mathcal{X}, \mathcal{Y}, \mathcal{C}_{\mathcal{XZ}} * \mathcal{C}_{\mathcal{ZY}})$ [7]. This definition of a cascade of two channels can be formalized in HOL4 as follows:

Definition 6 *(Cascade Channel).*

```
⊢ ∀X Z Y p f g. cascade_channel p X Z Y f g =
    channel p X Z f ∧
    channel p Z Y g ∧
    ∀x y. joint_distribution p X Y ({x},{y}) =
    ∑_z joint_distribution p X Z ({x}, {z}) *
    conditional_distribution p Y Z ({y}, {z})
```

5.2 Information Flow Analysis of Channels in Cascade

In order to analyze the information flow for the worst case scenario, i.e., when $\mathcal{A}$ recovers the critical information in one guess, we model the apriori distribution as a function of the maximum input distribution and the aposteriori behavior is expressed as a function of the maximum over $\mathcal{X}$ of the distribution of guessing a while observing b.

$$\textit{leakage} = \textit{Min-Entropy(X) - conditional Min-Entropy(X}|\,Y\textit{)}$$
$$IL_\infty(X,Y) = H_\infty(X) - H_\infty(X|Y)$$

Now, the leakage in a cascade of channels can be evaluated using Min-Entropy and the corresponding proof goal can be expressed in HOL4 as follows:

Theorem 9 *(Information Leakage of Channels in Cascade).*

Let $(\mathcal{X}, \mathcal{Y}, \mathcal{C}_{\mathcal{X}\mathcal{Y}})$ *be the cascade of* $(\mathcal{X}, \mathcal{Z}, \mathcal{C}_{\mathcal{X}\mathcal{Z}})$ *and* $(\mathcal{Z}, \mathcal{Y}, \mathcal{C}_{\mathcal{Z}\mathcal{Y}})$*. Then we have* $\mathcal{IL}_\infty(\mathcal{X}, \mathcal{Y}) \leq \mathcal{IL}_\infty(\mathcal{X}, \mathcal{Z})$ This theorem can be expressed in HOL4 as

```
⊢ ∀ p X Z Y f g.
   cascade_channel p X Z Y f g ∧
   FINITE (Ω) ∧
   Ω ≠ ∅ ∧
   events p = POW (Ω) ∧
   ∀x. 0 < distribution p Y {x} ∧
   ∀x. 0 < distribution p Z {x} ∧
   (∀x. x ∈ Ω ⇒ {x} ∈ events p) ⇒
   information_leakage p X Y ≤ information_leakage p X Z
```

Using some arithmetic simplification, the proof goal can be simplified to the level of vulnerabilities:

$$V_\infty(X|Y) \leq V_\infty(X|Z)$$

Now, using the property of cascade (*third conjunct in Definition 6*), we obtain

$$\begin{aligned} p(A = a|B = b) \quad &= \textstyle\sum_c p(A = a, C = c) * p(B = b|C = c) \\ &\leq \textstyle\sum_c max_a \quad p(A = a, C = c) * p(B = b|C = c) \end{aligned}$$

Next, we simplify the above subgoal by using the properties of summation along with the fact that the sum of the conditional distributions over the first state space of any random variable is equal to 1.

$$V(A \mid B) \leq \sum_c \max_a \quad p(A{=}a\ ,\ C{=}c)$$

The above subgoal can now be verified based on arithmetic simplification. This concludes the proof of Theorem 9, which consists of about 850 lines of HOL code.

5.3 Discussion

Due to the formal nature of the model and the soundness of the mechanical theorem prover, the analysis is guaranteed to be free of approximation and precision errors and thus the results obtained are mathematically precise and confirmed the results of paper-and-pencil based analysis approaches. This precision of analysis is a novelty that, to the best of our knowledge, has not been achieved by any other existing computer-based probabilistic analysis approaches. In the Definition 6 of the cascade channel behavior, the transition functions, f and g, are

general functions that provide generic results. In model checking approach parameters and functions should be specified. Furthermore the result verified in Theorem 9 can be extended to the Min-Entropy analysis of information leakage of n channels in cascade using induction techniques. We can prove that the Min-Entropy leakage of n channels in cascade will not exceed the leakage of the first channel. The main key to verify this property is the definition of the cascade condition. Mathematically, we can express the connection of n channels as follows

Let X_0 be the random variable modeling the input of the system and X_n the one modeling the output, thus

$$\forall \text{ i. } (0 \leq \text{i} \leq \text{n}) \Rightarrow \text{P}(X_0,X_i) = \sum_{\mathcal{X}_{i-1}} \text{P}(X_0,X_{i-1}) * \text{P}(X_i|X_{i-1})$$

Based on what we defined previously and what already existed, this condition can be formalized in HOL4 as

```
⊢ ∀X p f n. n_cascade_channel p X n f =
   ∀i. (1≤i≤n) ⇒ channel p (X (i-1)) (X i) (f i) ∧
   ∀x y i. joint_distribution p (X 0) (X i) (x,y) =
      ∑ joint_distribution p (X 0) (X (i-1)) (x,z) *
       z
      conditional_distribution p (X (i-1)) (X i) (z,y)
```

The ability to express and verify generic properties, quantified for all values of the variables, is the main strength of theorem proving as can be seen from the above definition and the property related to the information leakage of n channels in cascade. This property is an ongoing task, once verified, can hold for any number of cascade of channels and can be specialized to obtain expression and values for particular scenarios. Probabilistic model checking, which is the other main stream formal method, cannot provide such generic results due to the inherent state-space explosion problem.

6 Conclusion

This paper presents a formalization of vulnerability, belief-vulnerability, Min-Entropy and Belief Min-Entropy in higher-order logic. These metrics provide more reliable information flow analysis compared to the traditional definitions of quantitative information flow based on Shanon entropy for some corner cases. One such threat model being the case when an adversary can guess the secret input value in one try, given the observable output. The proposed formalization can be built upon to conduct the information flow analysis within the sound core of a theorem prover and thus the analysis is guaranteed to be free of approximation and precision errors. For illustration purposes, we performed the information flow analysis of a cascade of two channels using the HOL4 theorem prover and the analysis results were found to be generic and accurate.

The proposed higher-order-logic formalization can be used in analyzing many other applications. We are particularly aiming to apply it for the formal information flow analysis of the Crowds protocol [18] and Freenets [12]. Moreover, our work can be extended to analyze information flow in a reverse way, i.e. starting from a specific leakage bound we evaluate the input set with respect to the output set. This formalization can be used to formally ensure a specific level of security of critical information.

References

1. Andrea, S.: Possibilistic information theory: A coding theoretic approach. Fuzzy Sets Systems 132(1), 11–32 (2002)
2. Backes, M., Kopf, B., Rybalchenko, A.: Automatic discovery and quantification of information leaks. In: Proceedings IEEE Symposium on Security and Privacy, pp. 141–153. IEEE Computer Society (2009)
3. Coble, A.R.: Anonymity, information, and machine-assisted proof. Technical report, University of Cambridge, Computer Laboratory, Cambridge UK (July 2010)
4. Coble, A.R.: Anonymity, information, and machine-assisted proof. PhD thesis, King's College, University of Cambridge, Cambridge UK (2010)
5. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley-Interscience (1991)
6. Nguyen, H.T., Dubois, D., Prade, H.: Fundamentals of fuzzy sets, possibility theory, probability and fuzzy sets: Misunderstandings, bridges and gaps. In: Fundamentals of Fuzzy Sets. The handbooks of Fuzzy Sets Series, pp. 343–438. Kluwer (2000)
7. Espinoza, B., Smith, G.: Min-entropy leakage of channels in cascade. In: Barthe, G., Datta, A., Etalle, S. (eds.) FAST 2011. LNCS, vol. 7140, pp. 70–84. Springer, Heidelberg (2012)
8. Halpern, J.Y., O'Neill, K.R.: Anonymity and information hiding in multiagent systems. Journal of Computer Security 13(3), 483–514 (2005)
9. Hamadou, S., Sassone, V., Palamidessi, C.: Reconciling belief and vulnerability in information flow. In: Proceedings IEEE Symposium on Security and Privacy, pp. 79–92. IEEE Computer Society (2010)
10. Helali, G.: Formal analysis of information flow using min-entropy and belief min-entropy, `http://hvg.ece.concordia.ca/projects/prob-it/min_beliefInfo.php`
11. Hölzl, J.: Construction and Stochastic Applications of Measure Spaces in Higher-Order Logic. PhD thesis, Institut für Informatik, Technische Universität München, Germany (October 2012)
12. Clarke, I., Sandberg, O., Wiley, B., Hong, T.W.: Freenet: A distributed anonymous information storage and retrieval system. In: Federrath, H. (ed.) Designing Privacy Enhancing Technologies. LNCS, vol. 2009, pp. 46–66. Springer, Heidelberg (2001)
13. Palamidessi, C., Chatzikokolakis, K., Panangaden, P.: Anonymity Protocols as Noisy Channels. Information and Computation 206(2-4), 378–401 (2008)
14. Massey, J.L.: Guessing and entropy. In: Proceedings IEEE International Symposium on Information Theory, p. 204 (1994)
15. Mhamdi, T.: Information-Theoretic Analysis using Theorem Proving. PhD thesis, Department of Electrical and Computer Engineering, Concordia University (December 2012)

16. Mhamdi, T., Hasan, O., Tahar, S.: Formalization of entropy measures in HOL. In: van Eekelen, M., Geuvers, H., Schmaltz, J., Wiedijk, F. (eds.) ITP 2011. LNCS, vol. 6898, pp. 233–248. Springer, Heidelberg (2011)
17. Mhamdi, T., Hasan, O., Tahar, S.: Quantitative analysis of information flow using theorem proving. In: Aoki, T., Taguchi, K. (eds.) ICFEM 2012. LNCS, vol. 7635, pp. 119–134. Springer, Heidelberg (2012)
18. Reiter, M.K., Rubin, A.D.: Crowds: anonymity for web transactions. ACM Transactions on Information Systems Security 1(1), 66–92 (1998)
19. Renyi, A.: On measures of entropy and information. In: Proceedings Berkeley Symposium on Mathematics, Statistics and Probability, pp. 547–561 (1961)
20. Schneider, S., Sidiropoulos, A.: CSP and anonymity. In: Martella, G., Kurth, H., Montolivo, E., Bertino, E. (eds.) ESORICS 1996. LNCS, vol. 1146, pp. 198–218. Springer, Heidelberg (1996)
21. Smith, G.: Principles of secure information flow analysis. In: Malware Detection. Advances in Information Security, pp. 291–307. Springer (2007)
22. 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)
23. Smith, G.: Quantifying information flow using min-entropy. In: Quantitative Evaluation of SysTems, pp. 159–167 (2011)