A Precise Information Flow Measure from Imprecise Probabilities

Sari Haj Hussein[1]

[1]Department of Computer Science
Aalborg University

2012-06-21

Literature

- Clarkson, M.R.; Myers, A.C.; Schneider, F.B.; , "Belief in information flow," Computer Security Foundations, 2005. CSFW-18 2005. 18th IEEE Workshop , vol., no., pp. 31- 45, 20-22 June 2005
- Clarkson, M.R., Myers, A.C., Schneider, F.B. Quantifying information flow with beliefs (2009) Journal of Computer Security, 17 (5), pp. 655-701
- Hussein, Sari Haj; , "Refining a Quantitative Information Flow Metric," New Technologies, Mobility and Security (NTMS), 2012 5th International Conference on , vol., no., pp.1-7, 7-10 May 2012
- ...A Precise Information Flow Measure from Imprecise Probabilities

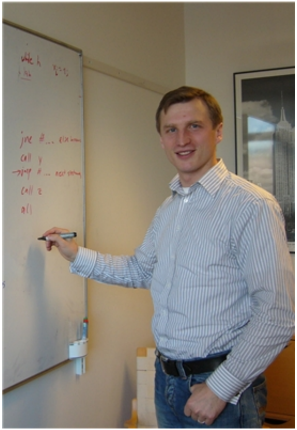
Literature

- Clarkson, M.R.; Myers, A.C.; Schneider, F.B.; , "Belief in information flow," Computer Security Foundations, 2005. CSFW-18 2005. 18th IEEE Workshop , vol., no., pp. 31- 45, 20-22 June 2005
- Clarkson, M.R., Myers, A.C., Schneider, F.B. Quantifying information flow with beliefs (2009) Journal of Computer Security, 17 (5), pp. 655-701
- Hussein, Sari Haj; , "Refining a Quantitative Information Flow Metric," New Technologies, Mobility and Security (NTMS), 2012 5th International Conference on , vol., no., pp.1-7, 7-10 May 2012
- ...A Precise Information Flow Measure from Imprecise Probabilities

Field

- Quantitative Information Flow (QIF) analysis
- Decide the number of bits that might be revealed from a program's secret input during the execution of that program

Qualitative...



Sabelfeld

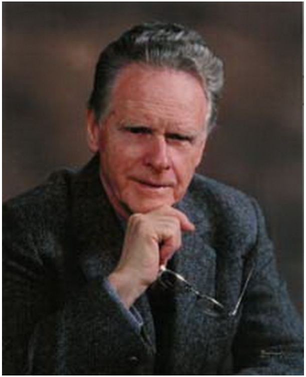
Problem

- The QIF metric by Clarkson et al
- It uses Bayesian inference
- It captures the improvement in the attacker's belief as she interacts with a program's execution
- Thereby it quantifies the flow

Contributions

- The paper presents a justified generalization of the analysis method done by Clarkson et al
- It highlights the weaknesses in the original work
- It shows that they are eliminated by way of the generalization
- The generalization is based on one of the theories of imprecise probabilities, namely the theory of evidence

Contributions



Dempster

Shafer

Frame of Discernment

Frame of Discernment (Sample Space)

- A set of possible worlds that an agent considers possible
- $\mathcal{W} = \{\text{password}, 123456, \text{qwerty}, \text{abc123}, \text{letmein}, 696969\}$

Closed-world Assumptions (Shafer's Model)

- Exclusiveness: At most one of the worlds in $\mathcal{W}$ is the true world
- Exhaustiveness: $\mathcal{W}$ contains all the possible worlds

Frame of Discernment

Frame of Discernment (Sample Space)

- A set of possible worlds that an agent considers possible
- $\mathcal{W} = \{\text{password}, 123456, \text{qwerty}, \text{abc123}, \text{letmein}, 696969\}$

Closed-world Assumptions (Shafer's Model)

- Exclusiveness: At most one of the worlds in $\mathcal{W}$ is the true world
- Exhaustiveness: $\mathcal{W}$ contains all the possible worlds

Frame of Discernment

Frame of Discernment (Sample Space)

- A set of possible worlds that an agent considers possible
- $\mathcal{W} = \{\text{password}, 123456, \text{qwerty}, \text{abc123}, \text{letmein}, 696969\}$

Closed-world Assumptions (Shafer's Model)

- ~~Exclusiveness:~~ At most one of the worlds in $\mathcal{W}$ is the true world
- Exhaustiveness: $\mathcal{W}$ contains all the possible worlds

Frame of Discernment

Dezert

Smarandache

Frame of Discernment

Example

- $\mathcal{PWC}$: if $p = g$ then $a := 1$ else $a := 0$ $\quad p \in \{A, B, C\}$
- $r = \{p, g, a\}$, $h = \{p\}$, $l = \{g, a\}$
- $\mathcal{W}_h = \prod_{X \in \{p\}} \mathcal{W}_X = \mathcal{W}_p = \{A, B, C\}$
- $$\mathcal{W}_l = \prod_{X \in \{g,a\}} \mathcal{W}_X = \mathcal{W}_g.\mathcal{W}_a = \{A, B, C\}.\{0, 1\} = \{(A, 0), (A, 1), (B, 0), (B, 1), (C, 0), (C, 1)\}$$
- $$\mathcal{W}_{h \cup l} = \prod_{X \in \{p,g,a\}} \mathcal{W}_X = \mathcal{W}_p.\mathcal{W}_g.\mathcal{W}_a = \{A, B, C\}.\{A, B, C\}.\{0, 1\} = \{(A, A, 0), \ldots\}$$

Frame of Discernment

Example

- $\mathcal{PWC}$: if $p = g$ then $a := 1$ else $a := 0$ $\quad p \in \{A, B, C\}$
- $r = \{p, g, a\}$, $h = \{p\}$, $l = \{g, a\}$
- $\mathcal{W}_h = \prod_{X \in \{p\}} \mathcal{W}_X = \mathcal{W}_p = \{A, B, C\}$
- $$\mathcal{W}_l = \prod_{X \in \{g,a\}} \mathcal{W}_X = \mathcal{W}_g . \mathcal{W}_a = \{A, B, C\}.\{0, 1\}$$ $$= \{(A, 0), (A, 1), (B, 0), (B, 1), (C, 0), (C, 1)\}$$
- $$\mathcal{W}_{h \cup l} = \prod_{X \in \{p,g,a\}} \mathcal{W}_X = \mathcal{W}_p . \mathcal{W}_g . \mathcal{W}_a = \{A, B, C\}.$$ $$= \{A, B, C\}.\{0, 1\} = \{(A, A, 0), \ldots\}$$

Belief Functions

- Frame of Discernment is too coarse
- Comparing the likelihood of worlds is not possible
- Belief functions is a numeric representation of uncertainty that enables full ordering of worlds

Belief Functions vs. Probability Measures

- The finite additivity property

$$Pro(X_1 \cup ... \cup X_n) = Pro(X_1) + ... + Pro(X_n)$$

- You are forced to work with singleton sets
- No overlapping sets

$$Pro(\{A, B\}) = 0.2,\ Pro(\{B, C\}) = 0.3$$

- No nested sets

$$Pro(\{A, B\}) = 0.2,\ Pro(\{A, B, C\}) = 0.3$$

Belief Functions vs. Probability Measures

- Ignorance is difficult to represent

$$Pro(\{A\}) = 0.2,\ Pro(\{A, B\}) = 0.0$$

- Contradiction is difficult to represent

$$Pro(\{A\}) = 0.2,\ Pro(\{B\}) = 0.3,\ Pro(\{\}) = 0.5$$

- Modeling collaboration is not possible
- Add the computational problem...

Mass Function

Mass Function

- $m : \mathcal{P}(\mathcal{W}_s) \to [0, 1]$
- $m(\emptyset) = 0, \sum_{A \in \mathcal{P}(\mathcal{W}_s)} m(A) = 1$
- $m(A)$ the degree of belief that the true world is in A

Belief Function

- $Bel : \mathcal{P}(\mathcal{W}_s) \to [0, 1]$
- $Bel(A) = \sum_{B \subseteq A} m(B)$

Mass Function

Mass Function

- $m : \mathcal{P}(\mathcal{W}_s) \rightarrow [0, 1]$
- $m(\emptyset) = 0, \sum_{A \in \mathcal{P}(\mathcal{W}_s)} m(A) = 1$
- $m(A)$ the degree of belief that the true world is in A

Belief Function

- $Bel : \mathcal{P}(\mathcal{W}_s) \rightarrow [0, 1]$
- $Bel(A) = \sum_{B \subseteq A} m(B)$

Belief Combination

Belief Combination

- Combining two pieces of evidence m_1 and m_2 from two independent sources
- $(m_1 \otimes m_2)(A) = k. \sum_{B \cap C = A} m_1(B).m_2(C)$
- $(m_1 \otimes m_2)(\emptyset) = 0,\ k^{-1} = \sum_{B \cap C \neq \emptyset} m_1(B).m_2(C)$
- $(m_1 \otimes m_2)(A) = k. \sum_{B \bowtie C = A} m_1(B).m_2(C)$
- $(m_1 \otimes m_2)(\emptyset) = 0,\ k^{-1} = \sum_{B \bowtie C \neq \emptyset} m_1(B).m_2(C)$

Belief Conditioning

Belief Conditioning

- Current agent's belief is captured using m
- A new piece of evidence that the true world is in B
- Agent can do a knowledge update...
- $m_B(A) = \begin{cases} k.\sum_{C \cap B = A} m(C) & \text{for } A \neq \emptyset \\ 0 & \text{for } A = \emptyset \end{cases}$
- $k^{-1} = \sum_{C \cap B \neq \emptyset} m(C)$

Belief Divergence

- We need to measure the divergence between 2 mass functions in an information-theoretic manner
- There is no out-of-the-box information-theoretic divergence in the theory of evidence
- Divergence measures, that are based on geometrical interpretations of mass functions, do not work
- We should derive a suitable divergence measure. How?
  1. Start with a divergence measure in probability theory
  2. Re-write this divergence in terms of information-theoretic functionals
  3. Generalize these functionals into the theory of evidence

Belief Divergence

- Kullback-Leibler $KL(p_1, p_2) = \sum_{x \in \mathcal{X}} p_1(x) \log \frac{p_1(x)}{p_2(x)}$
- Jensen-Shannon $JS(p_1, p_2) = 2S(\frac{p_1+p_2}{2}) - S(p_1) - S(p_2)$

Generalized Jensen-Shannon Divergence

- $GJS(m_1, m_2) = 2GS(\frac{m_1+m_2}{2}) - GS(m_1) - GS(m_2)$
- $GS(m) = AU(Bel) - GH(m)$

Belief Divergence

- Kullback-Leibler $KL(p_1, p_2) = \sum_{x \in \mathcal{X}} p_1(x) \log \frac{p_1(x)}{p_2(x)}$
- Jensen-Shannon $JS(p_1, p_2) = 2S(\frac{p_1+p_2}{2}) - S(p_1) - S(p_2)$

Generalized Jensen-Shannon Divergence

- $GJS(m_1, m_2) = 2GS(\frac{m_1+m_2}{2}) - GS(m_1) - GS(m_2)$
- $GS(m) = AU(Bel) - GH(m)$

In the paper...

- Imperative while-language
- Lift the syntax and semantics of it
- We are able to write program source code in terms of mass functions

In the paper...

- Start from an attacker's model
- Show how an attacker updates her knowledge from interacting with a program execution
- The arithmetic toolbox on beliefs and the execution rules of commands in the lifted language are extensively used here

In the paper...

- $\mathcal{PWC}$: if $p = g$ then $a := 1$ else $a := 0$ $\quad p \in \{A, B, C\}$

TABLE V
THE ATTACKER'S PREBELIEF AND POSTBELIEF IN EXPERIMENT 1

$\mathcal{P}(\mathcal{W}_h)$	m_{pre}	m'_{post}	m''_{post}
$\{A\}$	.98	1	0
$\{B, C\}$	.02	0	1

In the paper...

- $\mathcal{PWC}$: if $p = g$ then $a := 1$ else $a := 0$ $\quad p \in \{A, B, C\}$

TABLE VII

THE ATTACKER'S PREBELIEF AND POSTBELIEF IN EXPERIMENT 2

$\mathcal{P}(\mathcal{W}_h)$	m_{pre}	m'_{post}	m''_{post}
$\{A, B\}$	.98	0	0
$\{A, B, C\}$	.02	0	0
$\{A\}$	0	1	0
$\{B\}$	0	0	.98
$\{B, C\}$	0	0	.02

Measuring Information Flow

- When beliefs are involved then flow is the improvement in the accuracy of an attacker's belief

Our Flow Measure

- The accuracy of the attacker's prebelief is $GJS(m_{pre}, \dot{m}_h)$
- The accuracy of the attacker's postbelief is $GJS(m_{post}, \dot{m}_h)$
- $$\begin{aligned} Q &= GJS(m_{pre}, \dot{m}_h) - GJS(m_{post}, \dot{m}_h) \\ &= 2GS(\tfrac{m_{pre}+\dot{m}_h}{2}) - 2GS(\tfrac{m_{post}+\dot{m}_h}{2}) \\ &\quad - GS(m_{pre}) + GS(m_{post}) \end{aligned}$$

Measuring Information Flow

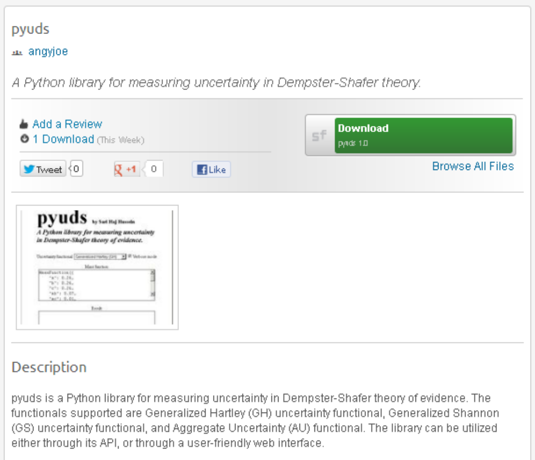
- When beliefs are involved then flow is the improvement in the accuracy of an attacker's belief

Our Flow Measure

- The accuracy of the attacker's prebelief is $GJS(m_{pre}, \dot{m}_h)$
- The accuracy of the attacker's postbelief is $GJS(m_{post}, \dot{m}_h)$
- $$\begin{aligned} \mathcal{Q} &= GJS(m_{pre}, \dot{m}_h) - GJS(m_{post}, \dot{m}_h) \\ &= 2GS(\tfrac{m_{pre}+\dot{m}_h}{2}) - 2GS(\tfrac{m_{post}+\dot{m}_h}{2}) \\ &\quad - GS(m_{pre}) + GS(m_{post}) \end{aligned}$$

In the paper...

- Sample flow calculations for the experiments



In the paper...

- The measure has the bounds $\varrho_{\mathcal{Q}} = [-\eta, \eta]$
- The space of the exhaustive search can be easily determined

Reflection and Future Work

- Probability theory has its base in set theory, but imprecise probabilities do not!
- The application of imprecise probabilities in fields other than QIF could be rewarding
- Subjective logic by Jøsang is good at trust modeling but does not work in QIF
- Set of stronger properties related to KL and JS whose proofs could be rewarding
- Could be interesting to do simulation using larger frames of passwords
- Could be interesting to look at guesswork in this setting

Pouly

Thank You!