# Refining a Quantitative Information Flow Metric

Sari Haj Hussein[1]

[1]Department of Computer Science
Aalborg University

2012-01-13

# Information Flow Analysis

- Information flow analysis aims at keeping track of a program's secret input during the execution of that program.

# Information Flow Analysis Techniques

- Qualitative techniques. prohibit flow from a program's secret input to its public output
  - Expensive or rarely satisfied by real programs
  - No distinguishment between acceptable and unacceptable flows
  - Conceptual and boring
- Quantitative techniques. establish limits on the number of *bits* that might be revealed from a program's secret input
  - Mainly based on information theory
  - More tangible

**Literature Observation**

Much work on qualitative, less on quantitative

# Information Flow Analysis Techniques

- Qualitative techniques. prohibit flow from a program's secret input to its public output
  - Expensive or rarely satisfied by real programs
  - No distinguishment between acceptable and unacceptable flows
  - Conceptual and boring
- Quantitative techniques. establish limits on the number of *bits* that might be revealed from a program's secret input
  - Mainly based on information theory
  - More tangible

**Literature Observation**

Much work on qualitative, less on quantitative

# Information Flow Analysis Techniques

- Qualitative techniques. prohibit flow from a program's secret input to its public output
  - Expensive or rarely satisfied by real programs
  - No distinguishment between acceptable and unacceptable flows
  - Conceptual and boring
- Quantitative techniques. establish limits on the number of *bits* that might be revealed from a program's secret input
  - Mainly based on information theory
  - More tangible

**Literature Observation**

Much work on qualitative, less on quantitative

# Problem Description

- The quantitative metric by Clarkson et al.
- It is the first to address attacker's belief in quantifying information flow
- This metric reports counter-intuitive flow quantities that are inconsistent with the size of a program's secret input.

# Problem Impact

- We cannot determine the space of the exhaustive search that should be carried out in order to reveal the residual part of a program's secret input

# Informal Reasoning

- There is a flaw in the design of the metric
- We need to spot the source of that flaw
- Then we need to fix it!

# Uncertainty-based Information Flow Analysis

## Uncertainty-based Information Flow Analysis *Denning*

- $\mathcal{U}$ attacker's pre-uncertainty
- $\mathcal{U}'$ attacker's post-uncertainty
- Flow = reduction in uncertainty
- $\mathcal{R} = \mathcal{U} - \mathcal{U}'$
- $\mathcal{R} \leq 0 \Rightarrow$ increase in uncertainty $\Rightarrow$ absence of flow
- $\mathcal{R} > 0 \Rightarrow$ decrease in uncertainty $\Rightarrow$ we have flow
- Notice that $\mathcal{R}$ ignores reality by measuring $\mathcal{U}$ and $\mathcal{U}'$ against each other, instead of against reality

# Plausible Range

- If attacker's belief is captured using a probability distribution, uncertainty is computed using Shannon uncertainty functional

**Shannon Uncertainty Functional**

- $X$ a discrete random variable with alphabet $\mathcal{X}$
- $p$ a probability distribution function on $X$
- $S(p) = -\sum_{x \in \mathcal{X}} p(x) \log p(x)$

- The range of $S$ is $[0, \log|\mathcal{X}|] \Rightarrow \varrho_{\mathcal{R}} = [-\log|\mathcal{X}|, \log|\mathcal{X}|]$
- This is plausible since $\log|\mathcal{X}|$ is the size of a program's secret input

# Plausible Range

- If attacker's belief is captured using a probability distribution, uncertainty is computed using Shannon uncertainty functional

## Shannon Uncertainty Functional

- $X$ a discrete random variable with alphabet $\mathcal{X}$
- $p$ a probability distribution function on $X$
- $S(p) = -\sum_{x \in \mathcal{X}} p(x) \log p(x)$

- The range of $S$ is $[0, \log|\mathcal{X}|] \Rightarrow \varrho_{\mathcal{R}} = [-\log|\mathcal{X}|, \log|\mathcal{X}|]$
- This is plausible since $\log|\mathcal{X}|$ is the size of a program's secret input

# Plausible Range

- If attacker's belief is captured using a probability distribution, uncertainty is computed using Shannon uncertainty functional

## Shannon Uncertainty Functional

- $X$ a discrete random variable with alphabet $\mathcal{X}$
- $p$ a probability distribution function on $X$
- $S(p) = -\sum_{x \in \mathcal{X}} p(x) \log p(x)$

- The range of $S$ is $[0, \log|\mathcal{X}|] \Rightarrow \varrho_{\mathcal{R}} = [-\log|\mathcal{X}|, \log|\mathcal{X}|]$
- This is plausible since $\log|\mathcal{X}|$ is the size of a program's secret input

# Size-consistent QIF Quantifier

## Size-consistent QIF Quantifier

- $QUAN$ a QIF quantifier
- $\eta$ the size of a program's secret input
- $QUAN$ is size-consistent if $\mathcal{QUAN}_{max} \leq \eta$ and $\mathcal{QUAN}_{min} \geq -\eta$

# Clarkson Observation

- $PWC$ : if $p = g$ then $a := 1$ else $a := 0$
- Password space is $W_p = \{A, B, C\}$ $\Rightarrow$ password size is $\log |W_p| = \log 3 = 1.5849$ bits
- The correct password (the reality) is $C$
- Attacker's prebelief $b_H = [(A : 0.98), (B : 0.01), (C : 0.01)]$
- Attacker (naturally) feeds $PWC$ with $g = A$ and gets $a = 0$
- Attacker's postbelief $b_H^{'} = [(A : 0), (B : 0.5), (C : 0.5)]$
- $\mathcal{R} = -0.8386$ bits $\Rightarrow$ absence of flow
- But $b_H^{'}$ is nearer to reality than $b_H$ $\Rightarrow$ attacker has learnt something $\Rightarrow$ we have flow

# Clarkson Conclusion

- Uncertainty-based analysis is inadequate if input distributions represent attacker's beliefs

# Accuracy-based Information Flow Analysis

## Accuracy-based Information Flow Analysis

- Respect reality by measuring $b_H$ and $b_H^{'}$ against it, instead of against each other only
- Reality is denoted as $\sigma_H$ (password is $C$)
- Certainty about reality is then $\dot{\sigma}_H$ (password is $C$ with a probability of 1)
- Accuracy of $b_H = D(b_H \rightarrow \dot{\sigma}_H)$
- Accuracy of $b_H^{'} = D(b_H^{\prime} \rightarrow \dot{\sigma}_H)$
- Flow = improvement in accuracy
- Clarkson metric $\mathcal{Q} = D(b_H \rightarrow \dot{\sigma}_H) - D(b_H^{\prime} \rightarrow \dot{\sigma}_H)$

# Clarkson Choice of $D$

- Clarkson chose Kullback-Leibler divergence
- $D(b \to b') = \sum_{\sigma \in \mathcal{W}_p} b'(\sigma) . \log \frac{b'(\sigma)}{b(\sigma)}$
- $\mathcal{Q} = D(b_H \to \dot{\sigma}_H) - D(b'_H \to \dot{\sigma}_H)$
- $\mathcal{Q} = \sum_{\sigma \in \mathcal{W}_p} \dot{\sigma}_H(\sigma) . \log \frac{\dot{\sigma}_H(\sigma)}{b_H(\sigma)} - \sum_{\sigma \in \mathcal{W}_p} \dot{\sigma}_H(\sigma) . \log \frac{\dot{\sigma}_H(\sigma)}{b'_H(\sigma)}$
- $\mathcal{Q} = -\log b_H(\sigma_H) + \log b'_H(\sigma_H)$

# Puzzling Result

- $\mathcal{Q} = -\log 0.01 + \log 0.5 = 6.6438 - 1 =$ 5.6438 bits
- But the plausible range is $\varrho_{\mathcal{R}} = [-\log 3, -\log 3] = [-1.5849, 1.5849]$
- $\mathcal{Q}$ is not a size-consistent QIF quantifier

## Clarkson Argument

- $b_H$ is more erroneous than a uniform belief ascribing $1/3$ probability to each password $A$, $B$, and $C$
- Therefore a larger amount of information is required to correct $b_H$
- If $b_H$ is uniform, the attacker would learn a total of $\log 3$ bits

# Our Arguments

- We have shown that Clarkson argument is valid for deterministic programs, but incomplete for probabilistic ones
- We have further shown that the range of $\mathcal{Q}$ is $\varrho_{\mathcal{Q}} = (-\infty, -\log b_H(\sigma_H)]$

# Replacing the Construct

## Original Construct

- $\mathcal{I}_{Dis}(\sigma) = \log \frac{b'(\sigma)}{b(\sigma)}$

## Proposed Construct

- $\mathcal{I}'_{Dis}(\sigma) = \log \frac{b'(\sigma)}{\frac{b'(\sigma)+b(\sigma)}{2}}$

## Replacement Effect

- $\mathcal{I}'_{Dis}(\sigma) \leq \frac{1}{2}\mathcal{I}_{Dis}(\sigma)$

# Replacing the Construct

## Original Construct

- $\mathcal{I}_{Dis}(\sigma) = \log \frac{b'(\sigma)}{b(\sigma)}$

## Proposed Construct

- $\mathcal{I}'_{Dis}(\sigma) = \log \frac{b'(\sigma)}{\frac{b'(\sigma)+b(\sigma)}{2}}$

## Replacement Effect

- $\mathcal{I}'_{Dis}(\sigma) \leq \frac{1}{2}\mathcal{I}_{Dis}(\sigma)$

# Replacing the Construct

## Original Construct

- $\mathcal{I}_{Dis}(\sigma) = \log \frac{b'(\sigma)}{b(\sigma)}$

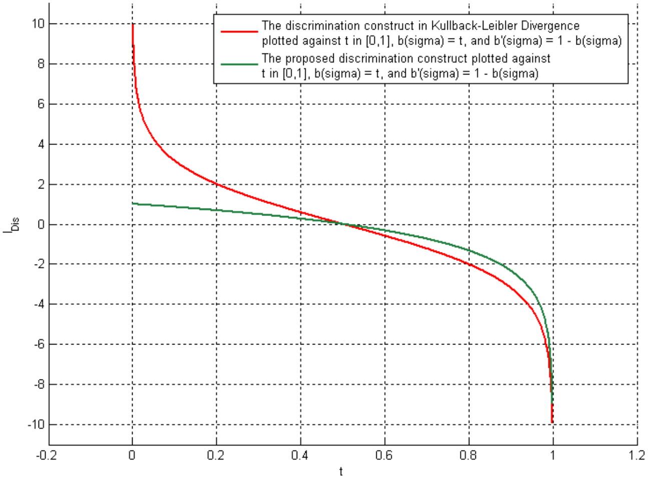
## Proposed Construct

- $\mathcal{I}'_{Dis}(\sigma) = \log \frac{b'(\sigma)}{\frac{b'(\sigma)+b(\sigma)}{2}}$

## Replacement Effect

- $\mathcal{I}'_{Dis}(\sigma) \leq \frac{1}{2}\mathcal{I}_{Dis}(\sigma)$

# Plot


The discrimination construct in Kullback-Leibler Divergence plotted against t in [0,1], b(sigma) = t, and b'(sigma) = 1 - b(sigma)
The proposed discrimination construct plotted against t in [0,1], b(sigma) = t, and b'(sigma) = 1 - b(sigma)
I_Dis
t
10
8
6
4
2
0
-2
-4
-6
-8
-10
-0.2
0
0.2
0.4
0.6
0.8
1
1.2

# Replacing the Divergence

## Original Divergence

- $D(b \to b') = \sum_{\sigma \in \mathcal{W}_p} b'(\sigma).\log \frac{b'(\sigma)}{b(\sigma)}$
- Average number of bits that are wasted by encoding events from a distribution $b'$ with a code based on a not-quite-right distribution $b$
- Information gain

## Proposed Divergence

- $D'(b \to b') = \sum_{\sigma \in \mathcal{W}_p} b'(\sigma).\log \frac{b'(\sigma)}{\frac{b'(\sigma)+b(\sigma)}{2}}$
- How much information is lost if we describe the two random variables that correspond to $b$ and $b'$ with their average distribution $(b'+b)/2$?
- Information radius

# Replacing the Divergence

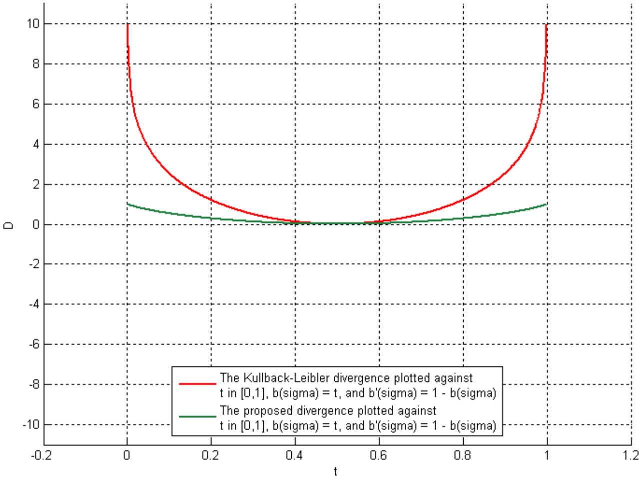
## Original Divergence

- $D(b \to b') = \sum_{\sigma \in \mathcal{W}_p} b'(\sigma).\log \frac{b'(\sigma)}{b(\sigma)}$
- Average number of bits that are wasted by encoding events from a distribution $b'$ with a code based on a not-quite-right distribution $b$
- Information gain

## Proposed Divergence

- $D'(b \to b') = \sum_{\sigma \in \mathcal{W}_p} b'(\sigma).\log \frac{b'(\sigma)}{\frac{b'(\sigma)+b(\sigma)}{2}}$
- How much information is lost if we describe the two random variables that correspond to $b$ and $b'$ with their average distribution $(b' + b)/2$?
- Information radius

# Plot



D
t
10
8
6
4
2
0
-2
-4
-6
-8
-10
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
The Kullback-Leibler divergence plotted against
t in [0,1], b(sigma) = t, and b'(sigma) = 1 - b(sigma)
The proposed divergence plotted against
t in [0,1], b(sigma) = t, and b'(sigma) = 1 - b(sigma)

# Refining to Normalization

## Normalized Metric

- $\mathcal{Q}' = D'(b_H \rightarrow \dot{\sigma}_H) - D'(b'_H \rightarrow \dot{\sigma}_H)$
- $\mathcal{Q}' = \sum_{\sigma \in \mathcal{W}_p} \dot{\sigma}_H(\sigma).\log \frac{\dot{\sigma}_H(\sigma)}{\frac{\dot{\sigma}_H(\sigma)+b_H(\sigma)}{2}} - \sum_{\sigma \in \mathcal{W}_p} \dot{\sigma}_H(\sigma).\log \frac{\dot{\sigma}_H(\sigma)}{\frac{\dot{\sigma}_H(\sigma)+b'_H(\sigma)}{2}}$
- $\mathcal{Q}' = \log \frac{2}{1+b_H(\sigma_H)} - \log \frac{2}{1+b'_H(\sigma_H)}$
- $\mathcal{Q}' = -\log(1 + b_H(\sigma_H)) + \log(1 + b'_H(\sigma_H))$

- We have shown that the range of $\mathcal{Q}'$ is $\varrho_{\mathcal{Q}'} = [-1, 1]$
- This does not make $\mathcal{Q}'$ size-consistent
- Nonetheless, $\varrho_{\mathcal{Q}'}$ is a plausible normalization (flow percentage) that is invariant with respect to the choice of the measurement unit

# Refining to Normalization

## Normalized Metric

- $\mathcal{Q}' = D'(b_H \rightarrow \dot{\sigma}_H) - D'(b'_H \rightarrow \dot{\sigma}_H)$
- $\mathcal{Q}' = \sum_{\sigma \in \mathcal{W}_p} \dot{\sigma}_H(\sigma).\log \frac{\dot{\sigma}_H(\sigma)}{\frac{\dot{\sigma}_H(\sigma)+b_H(\sigma)}{2}} - \sum_{\sigma \in \mathcal{W}_p} \dot{\sigma}_H(\sigma).\log \frac{\dot{\sigma}_H(\sigma)}{\frac{\dot{\sigma}_H(\sigma)+b'_H(\sigma)}{2}}$
- $\mathcal{Q}' = \log \frac{2}{1+b_H(\sigma_H)} - \log \frac{2}{1+b'_H(\sigma_H)}$
- $\mathcal{Q}' = -\log(1 + b_H(\sigma_H)) + \log(1 + b'_H(\sigma_H))$

- We have shown that the range of $\mathcal{Q}'$ is $\varrho_{\mathcal{Q}'} = [-1, 1]$
- This does not make $\mathcal{Q}'$ size-consistent
- Nonetheless, $\varrho_{\mathcal{Q}'}$ is a plausible normalization (flow percentage) that is invariant with respect to the choice of the measurement unit

# Refining to Actuality

## Actual Metric

- We want bit as the measurement unit
- Let $\eta$ be the size of a program's secret input in bits
- $\mathcal{Q}'' = \eta.\mathcal{Q}' = \eta.[-\log(1 + b_H(\sigma_H)) + \log(1 + b'_H(\sigma_H))]$

- We have shown that the range of $\mathcal{Q}''$ is $\varrho_{\mathcal{Q}''} = [-\eta.\log(1 + b_H(\sigma_H)), \eta.[1 - \log(1 + b_H(\sigma_H))]]$
- $\log(1 + b_H(\sigma_H)) \leq 1 \Rightarrow \mathcal{Q}''_{max} \leq \eta$ and $\mathcal{Q}''_{min} \geq -\eta \Rightarrow \mathcal{Q}''$ is size-consistent

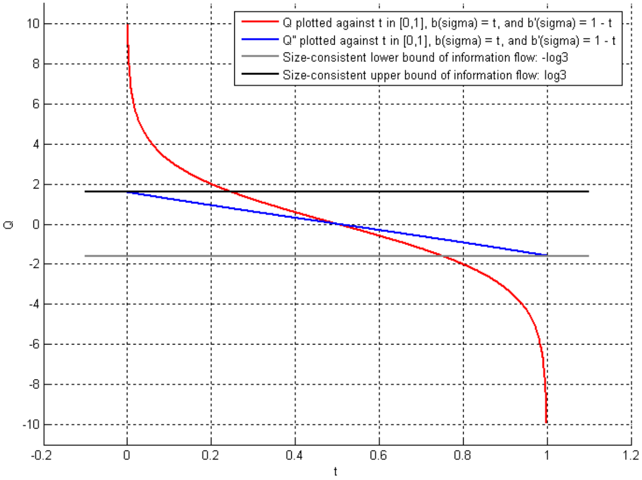
# Refining to Actuality

## Actual Metric

- We want bit as the measurement unit
- Let $\eta$ be the size of a program's secret input in bits
- $\mathcal{Q}'' = \eta.\mathcal{Q}' = \eta.[-\log(1 + b_H(\sigma_H)) + \log(1 + b'_H(\sigma_H))]$

- We have shown that the range of $\mathcal{Q}''$ is $\varrho_{\mathcal{Q}''} = [-\eta.\log(1 + b_H(\sigma_H)), \eta.[1 - \log(1 + b_H(\sigma_H))]]$
- $\log(1 + b_H(\sigma_H)) \leq 1 \Rightarrow \mathcal{Q}''_{max} \leq \eta$ and $\mathcal{Q}''_{min} \geq -\eta \Rightarrow \mathcal{Q}''$ is size-consistent

# Plot



Q plotted against t in [0,1], b(sigma) = t, and b'(sigma) = 1 - t
Q'' plotted against t in [0,1], b(sigma) = t, and b'(sigma) = 1 - t
Size-consistent lower bound of information flow: -log3
Size-consistent upper bound of information flow: log3
Q
t
10
8
6
4
2
0
-2
-4
-6
-8
-10
-0.2
0
0.2
0.4
0.6
0.8
1
1.2

# Interpreting the Refined Metric

- What does it mean to leak $k$ bits according to $\mathcal{Q}''$?
- $\mathcal{Q}'' = k$
- $\eta.[-\log(1 + b_H(\sigma_H)) + \log(1 + b'_H(\sigma_H))] = k$
- $\frac{\log(1+b'_H(\sigma_H))}{\log(1+b_H(\sigma_H))} = \frac{k}{\eta}$
- $\frac{1+b'_H(\sigma_H)}{1+b_H(\sigma_H)} = 2^{k/\eta}$
- $b'_H(\sigma_H) = 2^{k/\eta}.b_H(\sigma_H) + 2^{k/\eta} - 1$
- This corresponds to the increase in the likelihood of the attacker's correct guess

# Meaningfulness of the Bounds

- An informing flow equal to the upper bound of $\mathcal{Q}''$ is sufficient to make a fully uncertain attacker fully certain about the correct high state.
- $b_H(\sigma_H) = 0 \rightarrow \mathcal{Q}''_{max} = \eta.[1 - \log(1 + b_H(\sigma_H))] \rightarrow b'_H(\sigma_H) = 1$
- A misinforming flow equal to the lower bound of $\mathcal{Q}''$ is sufficient to make a fully certain attacker fully uncertain about the correct high state.
- $b_H(\sigma_H) = 1 \rightarrow \mathcal{Q}''_{min} = -\eta.\log(1 + b_H(\sigma_H)) \rightarrow b'_H(\sigma_H) = 0$

# Exhaustive Search Effort

- Assuming a program with a secret input of size $\eta$ bits.
- Assuming an informing flow of $k$ bits to an attacker
- $\mathcal{Q}''_{max} = \eta.[1 - \log(1 + b_H(\sigma_H))]$ tells us that $k \leq \eta$
- The space of the exhaustive search is $2^{\eta - k}$
- $\mathcal{Q}_{max} = -\log b_H(\sigma_H)$ tells us that $k > \eta$ is possible
- The exhaustive search space cannot be established, albeit that the secret input might have been partially revealed to the attacker

## Summary

- We presented a refinement of a QIF metric that bounds its reported results by a plausible range
- The results reported by the refined metric are easily associated with the exhaustive search effort
- We believe that the same can be done with other QIF quantifiers

# Thank You!