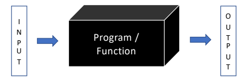
Quantitative Information Flow Analysis

CSCI 5271 Guest Lecture
Seonmo (Sean) Kim

Motivation



- An output has some data of an input.
- If the input contains some sensitive data, then output, too.
- The output should contain the intended amount of the input.
- An adversary wants to know the input by observing the output.

Motivation

- Consider two functions:

```
int numCheck(int input){
  if (input == 1234) {
    return 1;
  }
  return 0;
}
```

```
int numCheck2(int input){
  if (input mod 2 == 0) {
    return input;
  }
  return 1;
}
```

- The number of output values?

Motivation

- Consider two functions:

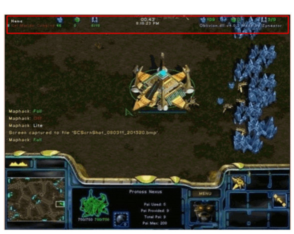
```
int numCheck(int input){
  if (input == 1234) {
    return 1;
  }
  return 0;
}
```

```
int numCheck2(int input){
  if (input mod 2 == 0) {
    return input;
  }
  return 1;
}
```

- The number of output values?
  - 2 vs $2^{31}+1$

Motivation

- There are many applications related to QIF analysis
- AI, games, financial programs, etc.
- Scalability



Quantitative Information Flow (QIF)

- Given a (deterministic or probabilistic) program P which takes a high input H and produces a low output L
- An adversary observes L and P may leak information from H (secret) to L (public)
- Measure the amount of information leaked about H

Early models of QIF

- Used the Shannon mutual information I(X;Y)
- Uncertainty
  - I(H; L) = H(H) – H(H | L)
  - information leaked = initial uncertainty – remaining uncertainty
    - the adversary's initial uncertainty before observing L
    - the adversary's remaining uncertainty after observing L
  - H(H) - I(H; L) = H(H | L)

Shannon entropy: initial uncertainty

- $H(X) = -\sum_{x \in \mathcal{X}} \Pr[\mathrm{X}{=}x] \cdot \log_2 \Pr[\mathrm{X}{=}x]$
- If H is a 32-bit integer and L := H
  - $\Pr[\mathrm{H}{=}x] = 1/2^{32}$, $\log_2 \Pr[\mathrm{H}{=}x] = \log_2 2^{-32} = -32$
  - $H(H) = -2^{32}(1/2^{32})(-32) = 32$

Shannon entropy: information leaked

- $I(X;Y) = H(X) - H(X \mid Y) = H(X) + H(Y) - H(X,Y)$
  - If X is determined by Y, then H(X|Y)=0.
- $I(H;L) = I(L;H) = H(L) - H(L \mid H) = H(L)$
- If H is a 32-bit integer and L := H
  - $I(H;L) = H(L) = H(H) = 32$
  - $\Pr[\mathrm{H}{=}x] = 1/2^{32}$, $\log_2 \Pr[\mathrm{H}{=}x] = \log_2 2^{-32} = -32$
  - $H(H) = -\sum_{x \in \mathcal{X}} \Pr[\mathrm{X}{=}x] \cdot \log_2 \Pr[\mathrm{X}{=}x] = -2^{32}(1/2^{32})(-32) = 32$
  - Remaining uncertainty: $H(H|L) = 32 - 32 = 0$

Shannon entropy

- $H(X) = -\sum_{x \in \mathcal{X}} \Pr[\mathrm{X}{=}x] \cdot \log_2 \Pr[\mathrm{X}{=}x]$
  - If H is a 32-bit integer and L := H, H(H) = 32
- $H(X \mid Y) = H(X) - I(X;Y)$
  - If H is a 32-bit integer and L := H, H(H | L) = 0
- $I(X;Y) = I(Y;X) = H(Y) - H(Y \mid X) = H(Y)$, if Y is determined by X
  - If H is a 32-bit integer and L := H, I(H ; L) = 32
- Exercise

Program	H (H)	I (H ; L)	H (H | L)
L := 0			
L := H & 0x0000ffff			

Shannon entropy

- $H(X) = -\sum_{x \in \mathcal{X}} \Pr[\mathrm{X}{=}x] \cdot \log_2 \Pr[\mathrm{X}{=}x]$
  - If H is a 32-bit integer and L := H, H(H) = 32
- $H(X \mid Y) = H(X) - I(X;Y)$
  - If H is a 32-bit integer and L := H, H(H | L) = 0
- $I(X;Y) = I(Y;X) = H(Y) - H(Y \mid X) = H(Y)$, if Y is determined by X
  - If H is a 32-bit integer and L := H, I(H ; L) = 32
- Exercise

Program	H (H)	I (H ; L)	H (H | L)
L := 0	32	0	32
L := H & 0x0000ffff	32	16	16

Alternative measurement

- Consider two programs:
  - if H mod 8 == 0 then L := H
    else L := 1
    - An adversary can guess H with probability 1/8
    - $P[L = 1] = 7/8$, $P[L = 8^n] = 1/2^{32}$ where $0 \le n < 29$
    - $I(H;L) = H(L) = \frac{7}{8}\log\frac{8}{7} + \frac{2^{29}}{2^{32}}\log 2^{32} \approx 0.169 + 4$
  - L := H & 0x0000001f
    - An adversary can guess H with probability $1/2^{27}$
    - $I(H;L) = H(L) = 5$
- Which one is more secure?

Alternative measurement

- Vulnerability
  - $V(X) = max_{x \in \mathcal{X}} \Pr[\mathrm{X}=x]$
- min-entropy
  - $H_\infty(X) = -\log_2 V(X)$
  - $H_\infty(X|Y) = -\log_2 V(X|Y)$
- information leaked = $H_\infty(H) - H_\infty(H|L)$
  - Let |X| be the number of possible values of X
  - $V(H) = \frac{1}{|H|}, V(H|L) = \frac{|L|}{|H|}$
  - $H_\infty(H) - H_\infty(H|L) = \log_2|H| - \log_2(|H|/|L|) = \log_2|L|$

Alternative measurement

- Consider two programs:
  - if H mod 8 == 0 then L := H
    else L := 1
    - $|L| = 2^{32-3} + 1$
    - Information Leakage = $\log_2|L| \approx 29$
  - L := H & 0x0000001f
    - $|L| = 2^5$
    - Information Leakage = $\log_2|L| = 5$

Applications

- Image anonymization and Kbattleship (PLDI 2008)
  - Computing a maximum flow of information
- Error reporting system (ASPLOS 2008)
- Heartbleed (VMCAI 2018)
  - Using the model counting technique to measure the leakage

Image Anonymization

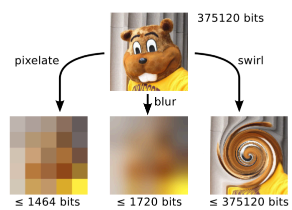
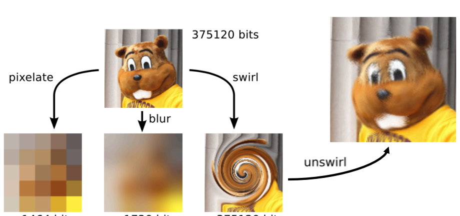
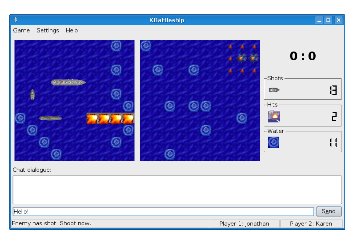


Image Anonymization

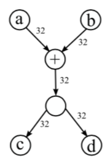


KBattleship



Flowcheck

- Dynamic analysis tool to measure an upper-bound estimate of the amount of information leaked
- Dynamic tainting
- Static control-flow regions
- c = d = a + b

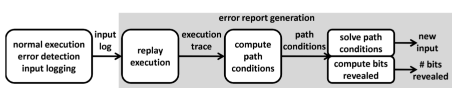


Error Reporting System

- Scenario

Error Reporting System

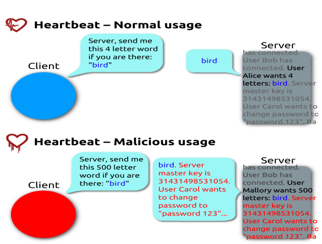
- Symbolic Execution
  - Generates path conditions based on symbolic or concrete inputs



Measuring privacy loss

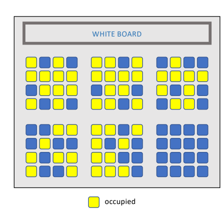
- For each condition (*op f*(.) *g*(.)), compute a *summary* for *f* and *g*
- Use a set of rules to compute the bound given the summaries
- Example
  - (add (bitwise-and x 1) 3)
    - (bitwise-and x 1) -> 0 or 1
    - (add (bitwise-and x 1) 3) -> 3 or 4

Heartbleed



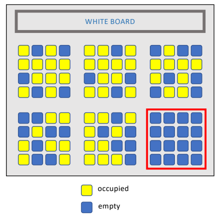
Exact Model Counting

- Brute-force counting
  - Go through every seat
  - Simple, but hard to scale



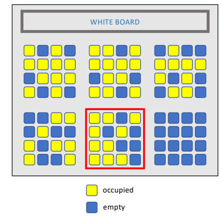
Exact Model Counting

- Brute-force counting
  - Go through every seat
  - Simple, but hard to scale
- DPLL-style counting
  - Detect a region that is empty
  - Faster, but still accounts for every seat



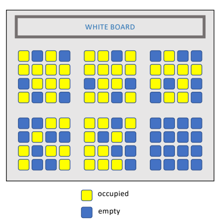
Approximate model counting

- Random sampling
  - Randomly pick a region
  - Count the number and scale up



Approximate model counting

- Random sampling
  - Randomly pick a region
  - Count the number and scale up
- Random hashing(AAAI 2006)
  - Everyone flips a coin k times
  - Leave if a tail is ever shown
  - Count the persons n
  - Approximately $2^k \cdot n$ persons



Q & A

Thank You:)