CANINE

A NetFlows Conversion/Anonymization Tool for Format Interoperability and Secure Sharing

Katherine Luo*, Yifan Li, Adam Slagell, William Yurick

SIFT Research Group

National Center for Supercomputing Applications (NCSA)

University of Illinois at Urbana-Champaign

FloCon05, Sep. 20, 2005

Motivations

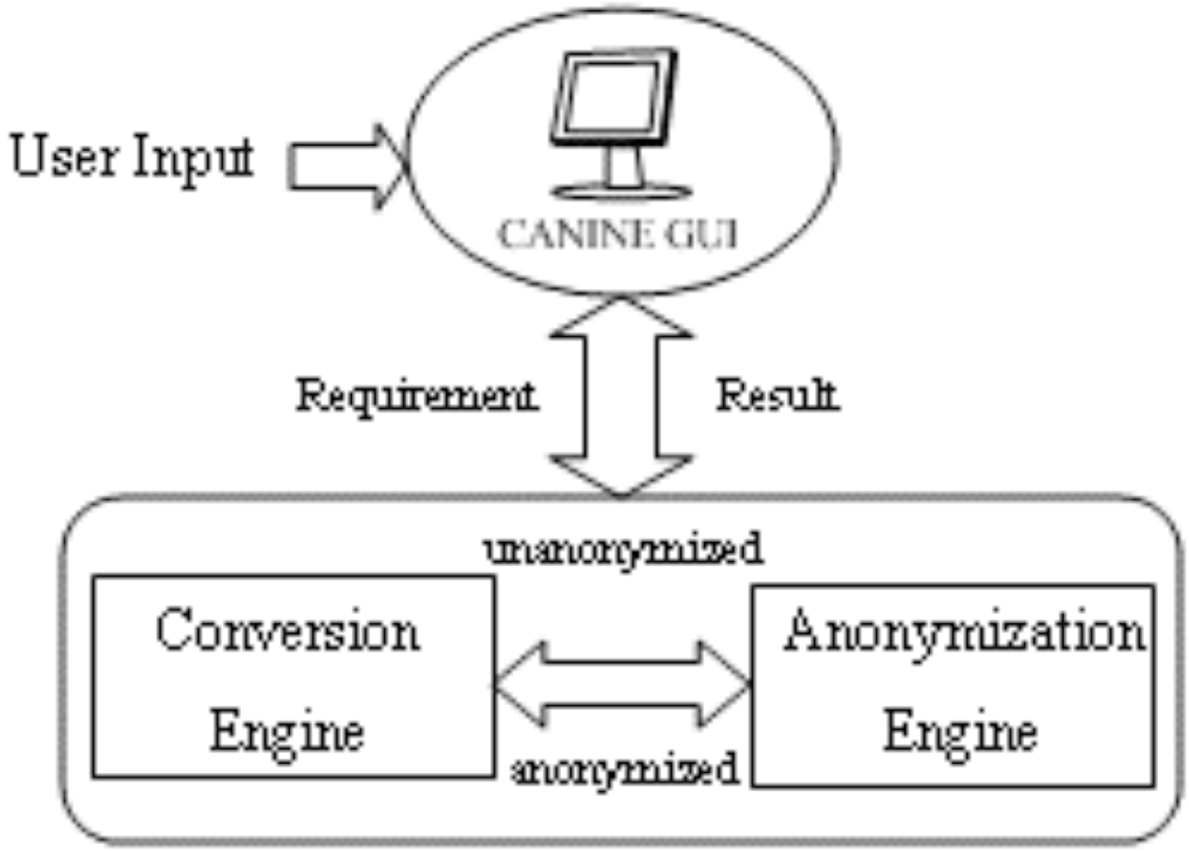
- NetFlows in multiple, incompatible formats
  - Network security monitoring tools usually support one or two NetFlows format
  - Need conversion of NetFlows between different formats
- Sensitive network information hinders log sharing
  - Log sharing necessary for research and study
  - Need anonymization of sensitive data fields

Our Solution: CANINE Tool

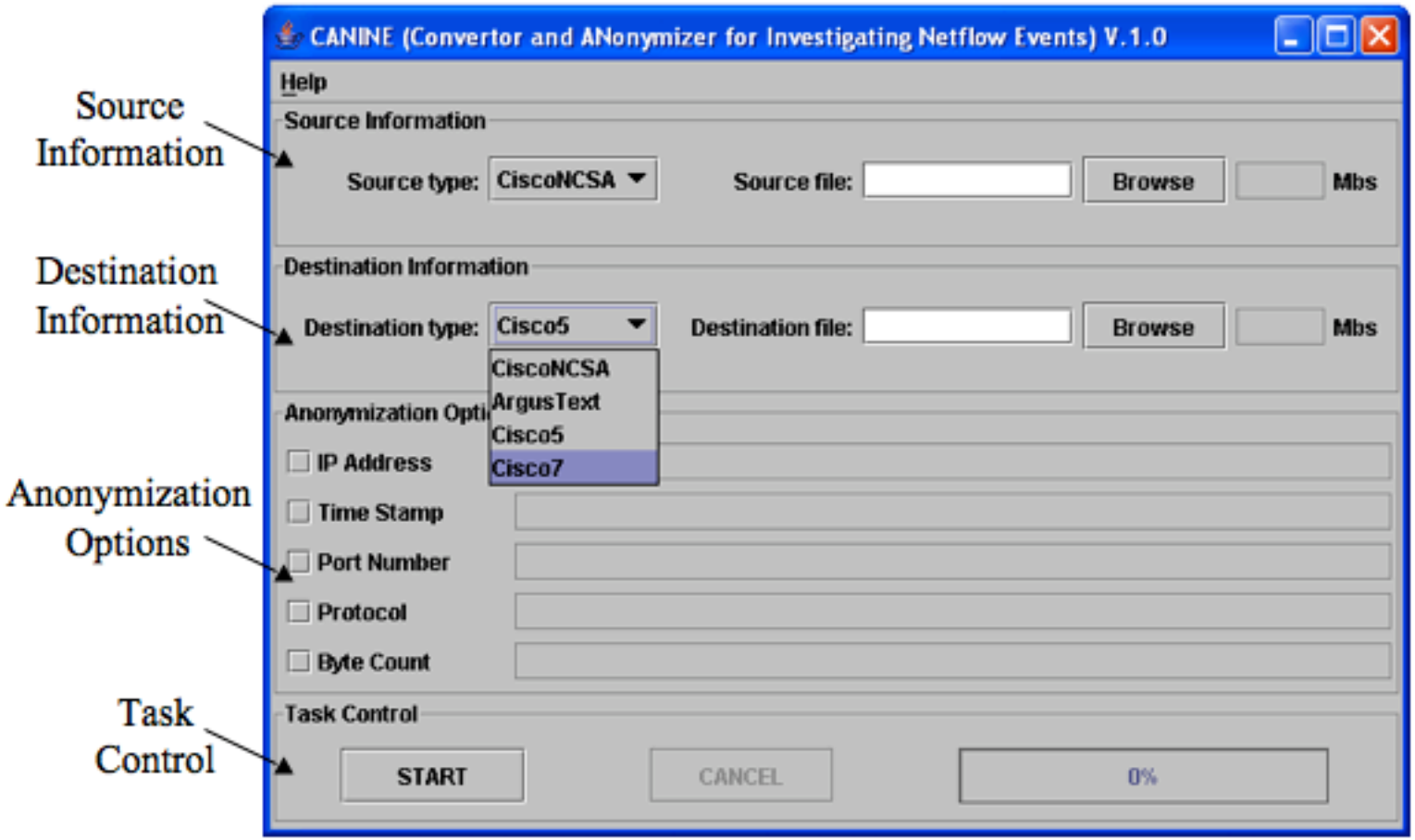
- CANINE: Converter and ANonymizer for Investigating Netflow Events
- Handles several NetFlow formats
  - Cisco V5 & V7, ArgusNCSA, CiscoNCSA, NFDump
- Anonymizes 5 types of data fields
  - IP, Timestamp, Port, Protocol and Byte Count

- Multiple anonymization levels
  - Various anonymization methods for some data field

System Architecture of CANINE



Main GUI of CANINE



Conversion & Anonymization Engine

- Conversion Engine
  - Parse the input NetFlow record into component data fields before anonymization
  - Reassemble the anonymized data component to desired NetFlow format
- Anonymization Engine
  - Contain a collection of anonymization algorithms
  - Anonymize data fields with designated methods

IP Address Anonymization

- Truncation
  - Zeroing out any number of LSBs
- Random Permutation
  - Generate a random IP number seeded by user input
- Prefix-preserving Pseudonymization
  - Match on n-bit prefix, based on Crypto-PAn

IP Address	Truncation (16-bit)	Random Permutation	Prefix-preserving
141.142.96.167	141.142.0.0	124.12.132.37	12.131.102.67
141.142.96.18	141.142.0.0	231.45.36.167	12.131.102.197
141.142.132.37	141.142.0.0	12.72.8.5	12.131.201.29

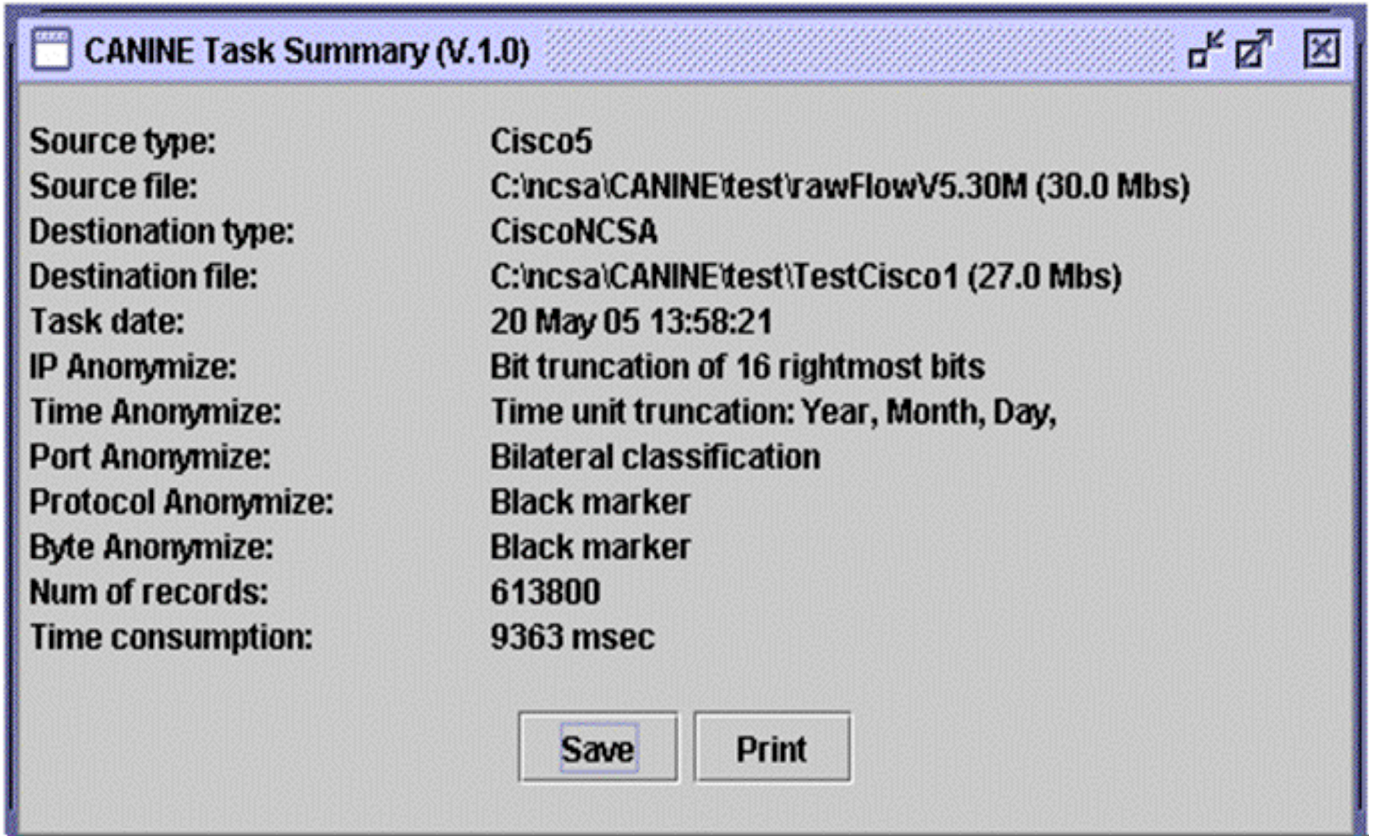
Timestamp Anonymization

- Time Unit Annihilation
  - Zeroing-out indicated subset of time units on end time
  - Start time is adjusted to keep the duration unchanged
- Random Time Shift
  - Pick a range for generating random shift
  - Shift all timestamps by the same amount
- Enumeration
  - Local sorting performs based on end time
  - Set the slide window size
  - Records sorted and equidistantly spaced

Port Number, Protocol, Byte Count Anonymization

- Port Number Anonymization
  - Bilateral classification
    - Replace with 0 or 65535 (the port smaller or larger than 1024)
  - Black marker
    - Replace with 0
- Protocol Anonymization
  - Black Maker
    - Replace with 255 (IANA reserved but unused number)
- Byte Count Anonymization
  - Black Marker
    - Replace with 0 (Impossible value in practice)

Task Summary Dialog



Summary and Future Work

- CANINE addressed two problems
  - Convert and anonymize NetFlow logs
  - Unique due to multiple anonymization levels
- Modifications on CANINE
  - Config file alternative to GUI
  - Streaming mode processing
- Research on multiple levels of anonymization scheme
  - Utility of the anonymized log
  - Security of the anonymization schemes

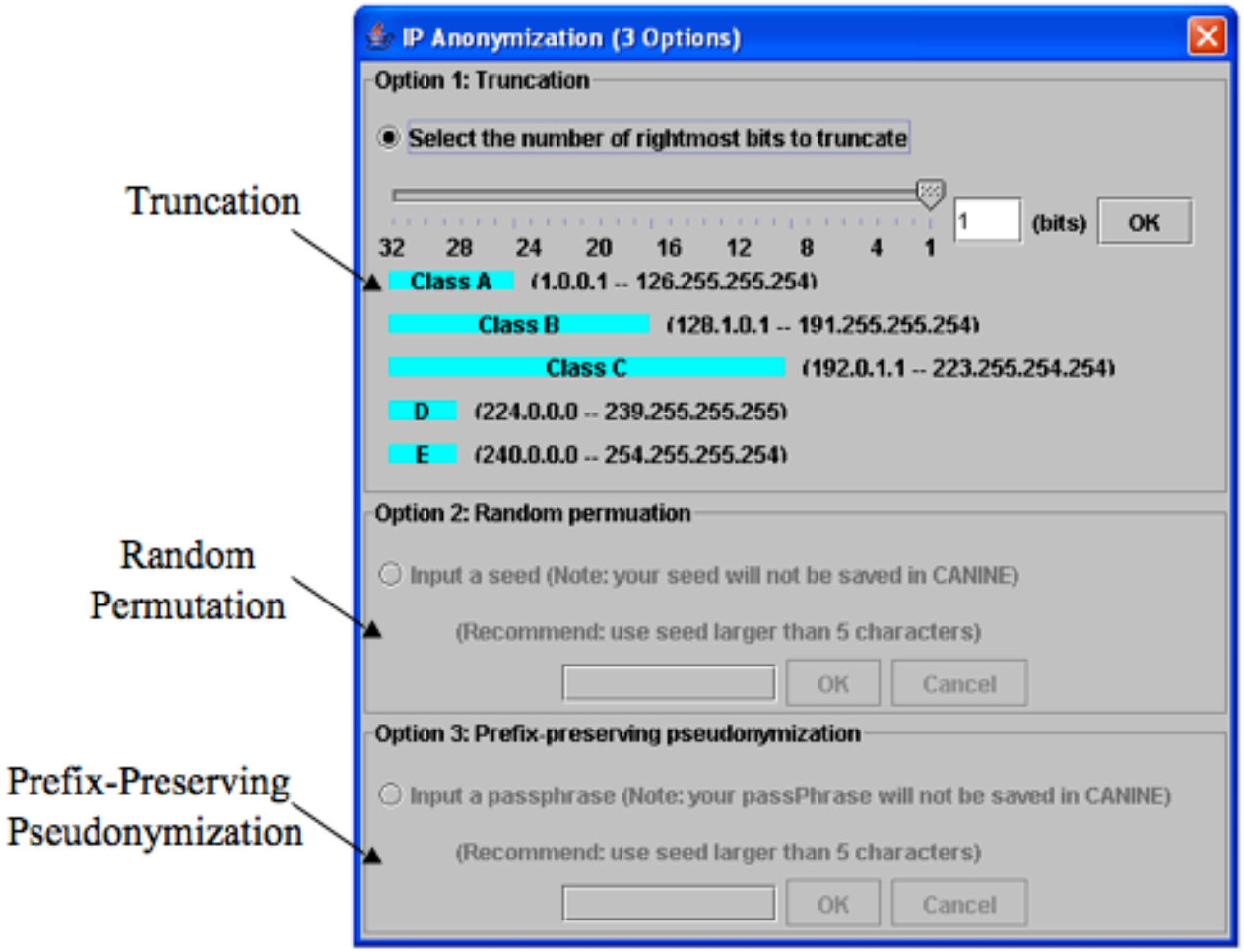
Download CANINE at

http://security.ncsa.uiuc.edu/distribution/CanineDownLoad.html

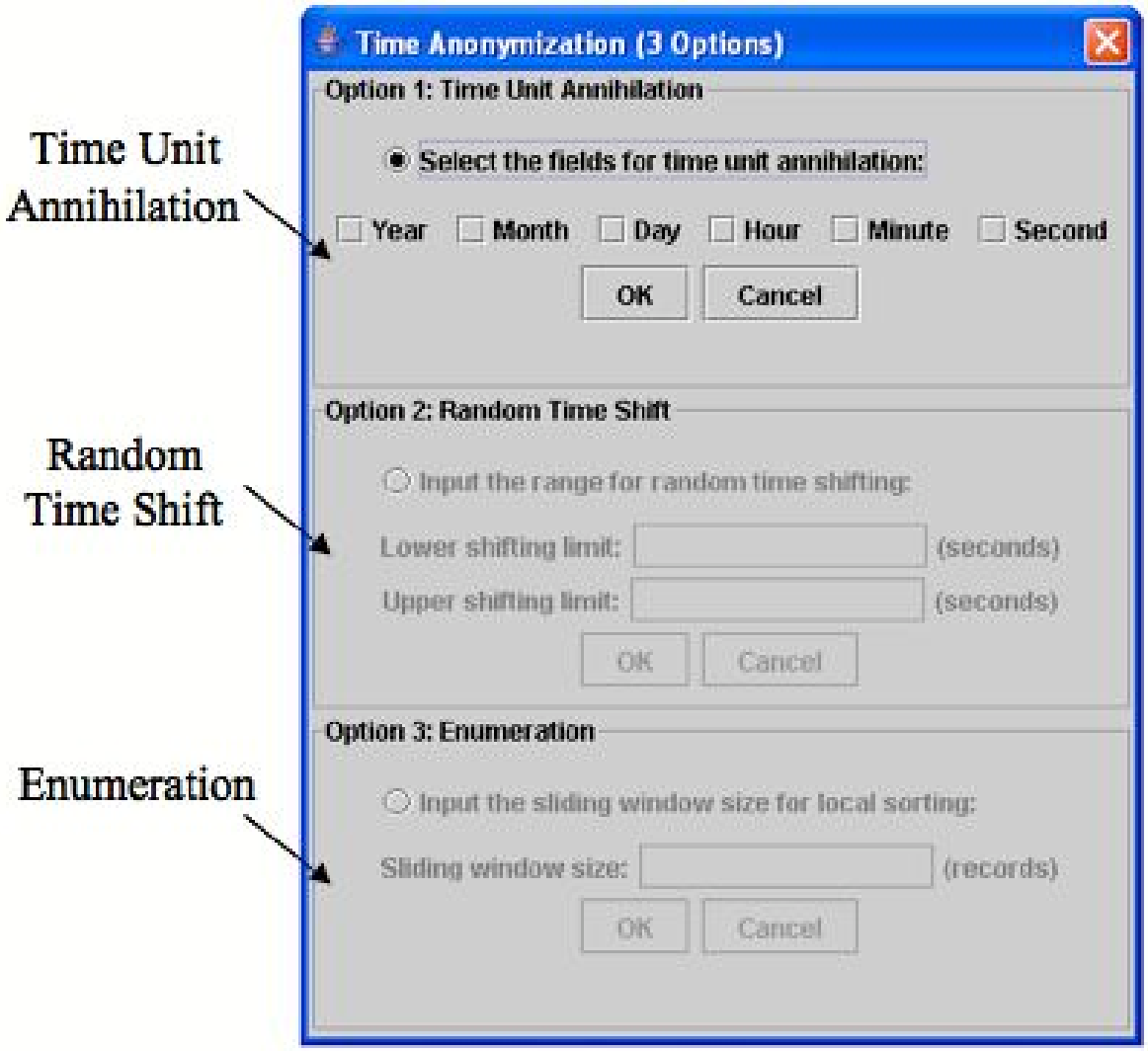
Thank you!

Questions?

IP Address Anonymization



Timestamp Anonymization



Port Number Anonymization

- Bilateral classification
  - –Decide the port is ephemeral or not
- Black marker