SWINBURNE UNIVERSITY OF TECHNOLOGY

Covert Channels and Machine Learning Traffic Classification

Sebastian Zander

szander@swin.edu.au

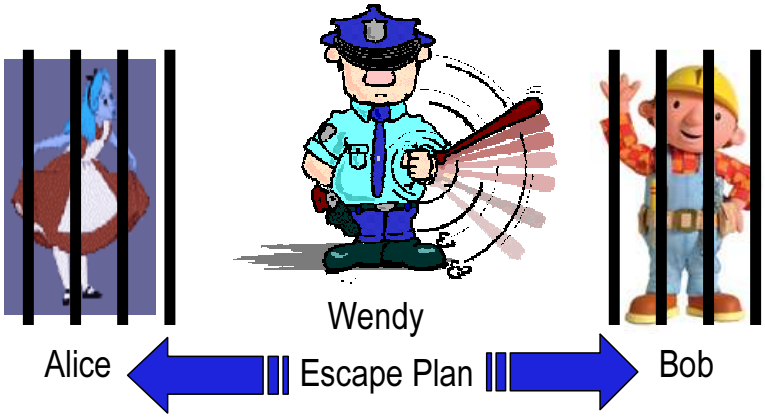
Centre for Advanced Internet Architectures (CAIA)
Swinburne University of Technology

Who am I?

- Research Fellow at CAIA
- Previously
  - Scientist at Fraunhofer FOKUS (Germany)
  - PhD student at CAIA
- Research interests: traffic classification, covert channels, network performance measurements
- Co-author of 30+ peer-reviewed papers, 2 IETF RFCs
- Collaborators: Grenville Armitage, Philip Branch, Steven Murdoch (Uni. Cambridge), Tanja Zseby (Fraunhofer FOKUS), Georg Carle (Tech. Uni. Munich), Benoit Claise (Cisco), ...

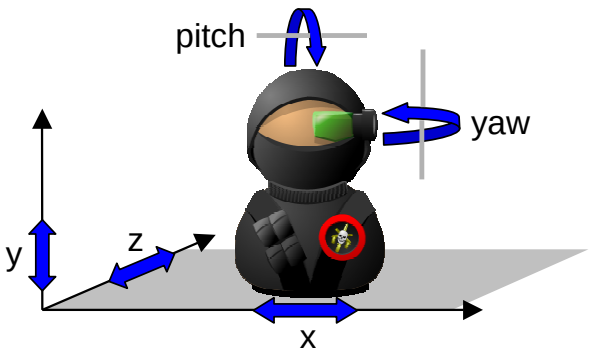
Covert Channels

- Encryption protects content of communication
- Existence of communication enough to take actions
- Covert channels hide existence of communication
- Potential users
  - Gov. agencies or criminals hiding communication
  - Hackers or spies ex-filtrating data
  - Users circumventing censorship or bypassing firewalls



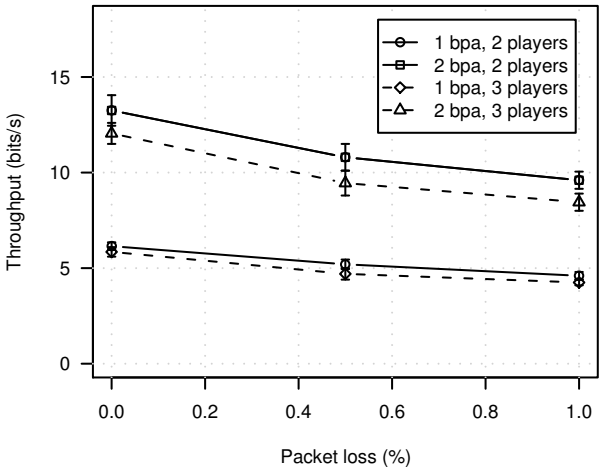
Noisy Covert Channels

- Many simple noise-free channels, e.g. unused bits...
  - Easy to detect and eliminate by security systems
- More complex channels are noisy
- Selected channels: IP Time-to-Live field, packet timing, games / virtual worlds, temperature-based
- Investigate noise characteristics, capacity, encoding schemes, countermeasures



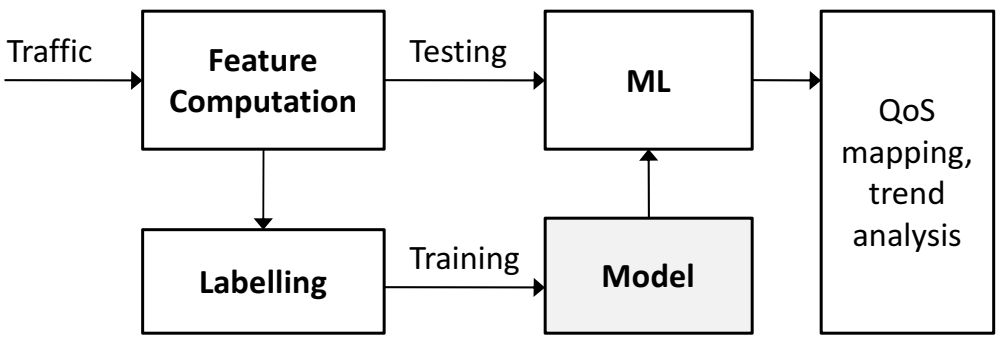
Covert Channels Findings

- Bit error rates up to 1–10% and bit sync. errors
- Capacity up to a few hundreds bits/second
- Sufficient for sending short messages, documents
- Some channels easy to detect but hard to eliminate and vice versa
- Machine Learning detects some channels with $>$ 95%



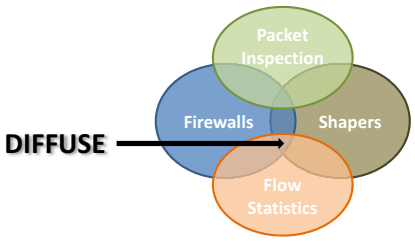
Machine-Learning Traffic Classification

- Port numbers unreliable (NATs, proxies, obfuscation)
- Payload inspection is limited (encryption, privacy)
  - Traffic features, e. g. packet length
  - Machine Learning (ML) feature-based classification
- ML too costly for low-end home DSL/Cable routers
  - Separate flow classification and treatment



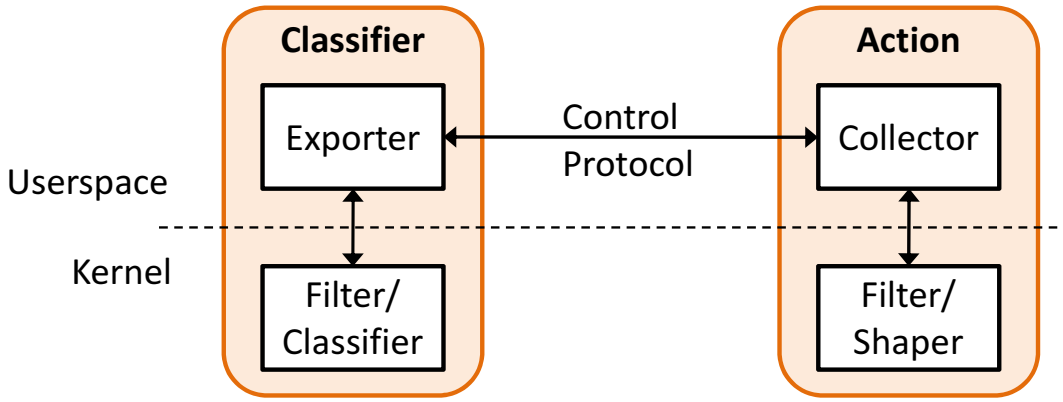
DIFFUSE Project

- DIstributed Firewall and Flow-shaper Using Statistical Evidence (DIFFUSE)
- Funded by Cisco University Research Program
- 12 month project started in June 2010
- Successor of 2005/2006 URP project
- Develop open-source prototype based on existing IPFW firewall
- Analyse classifier speed, accuracy, stability



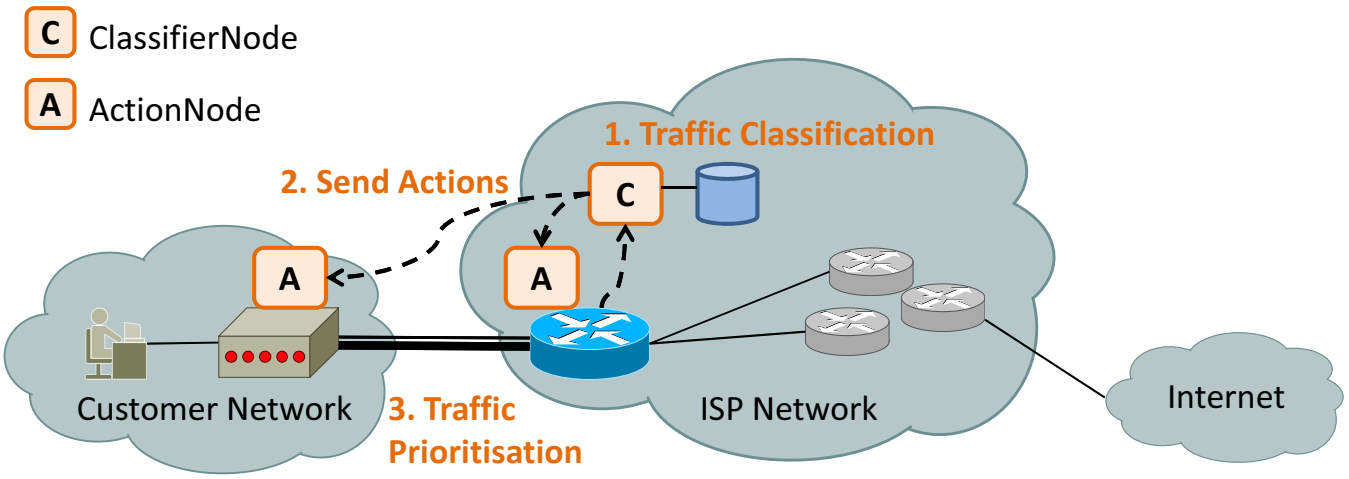
DIFFUSE Architecture

- Classifier Nodes (CNs) classify flows
- Action Nodes (ANs) treat flows based on class
- CNs control ANs via control protocol
- Extended rule language used to configure CNs, ANs



DIFFUSE Applications

- Automated QoS provision for interactive traffic
- Lawful Interception
- Detection and blocking malicious traffic



The End

Questions???