

# Network Traffic Classification

## Naive Bayes Classification

Kefei Lu

Department of Electrical and Computer Engineering  
University of Miami

2009-11-24

# The System Flow Chart

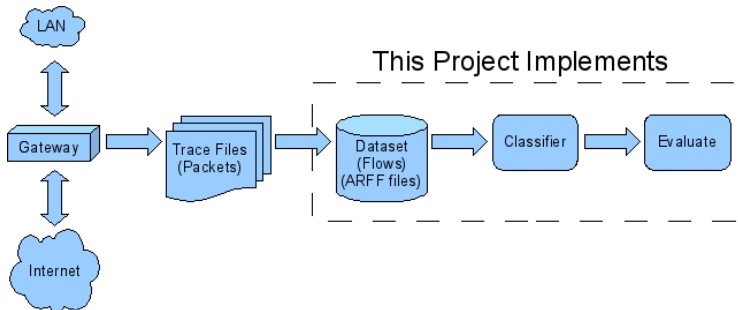


Figure: The flow chart of network traffic classification.

# From Packets To Flows

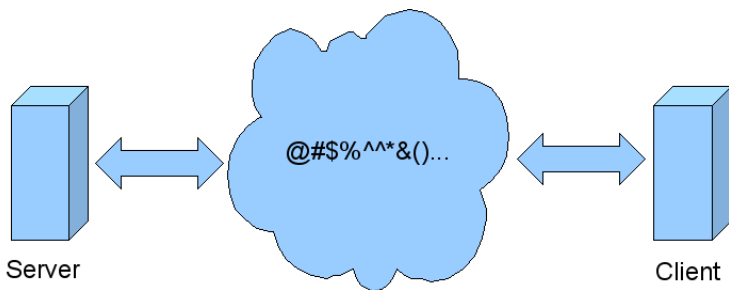


Figure: Internet Traffic (Flows)

- A flow consists of consecutive packets.
- A flow can be either directional/bidirectional
- A flow can be complete or incomplete
- Just use tools to identify flows!
- TCP Trace (a trace file parser)

# The Implemented Structure

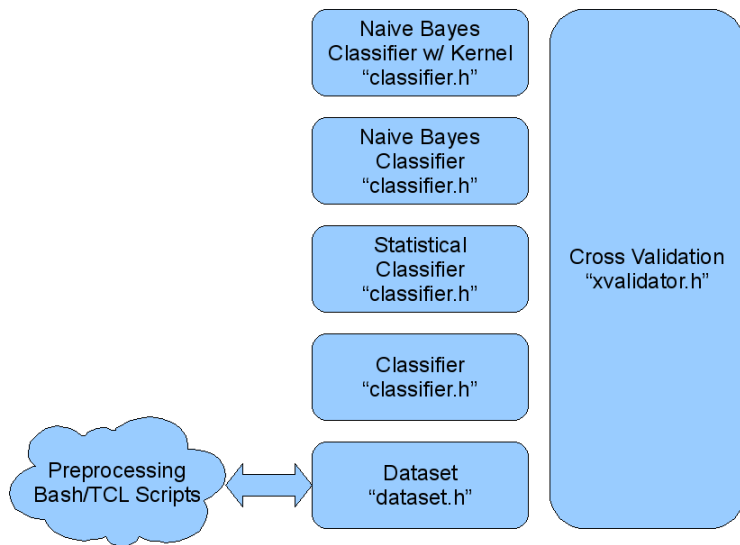


Figure: The Implemented Structure

# The Naive Bayes Method

The Maximum A Posteriori (MAP) criteria

$$i = \operatorname{argmax}_{i \in I} Pr(c_i | \mathbf{o}), \quad (1)$$

The Naive Bayes Method And The Assumptions

$$Pr(c_i | \mathbf{o}) = \frac{Pr(\mathbf{o} | c_i) Pr(c_i)}{Pr(\mathbf{o})} \quad (2)$$

$$= \frac{Pr(c_i) \prod_{k=0}^M Pr(a_k | c_i)}{\sum_{j=0}^N Pr(c_j) \prod_{k=0}^M Pr(a_k | c_j)}, \quad (3)$$

Welcome! It's on the web!  
Naive Bayes Classifier for Internet Traffic  
NBC4IT: <http://code.google.com/p/nbc4it/>