Profiling Internet Backbone Traffic: Behavior Models and Applications

PI: Prof. Zhi-Li Zhang (zhzhang@cs.umn.edu)
Student: Kuai Xu (kxu@cs.umn.edu)
Collaborator: Supratik Bhattacharyya

Why profile traffic?
• Changes in Internet traffic dynamics
  – increase in unwanted traffic, e.g., DDoS, worms
  – emergence of disruptive applications, e.g., P2P
  – new services on traditional ports, e.g., Skype
  – traditional service on non-standard ports
• Existing tools
  – rely on ports for identifying or classifying traffic
  – report volume-based heavy hitters
  – look for specific or known patterns
• Need better techniques to discover behavior patterns
  – help network operators secure and manage networks

Communication patterns

Extract significant clusters

Data pre-processing
• Aggregate packet streams into 5-tuple flows
• Group flows associated with same end hosts/ports into clusters

Behavior classifications

Case Studies

• Identify interesting events using typical profiles
  – server profiles on high ports, e.g., 60638
  – p2p traffic on alternative ports
  – exploit activities on unknown ports, e.g., an end host probing random dstIPs on dstPort 12347
• Rare behaviors
  – behavior patterns that rare happen are interesting
    – case study: exploit traffic from NAT boxes
• Deviant behaviors
  – clusters change from its usual BCs to a different
    – case study: a web server under DoS attack

Structural modeling

Conclusions

• Develop a systematic methodology to automatically discover and interpret communication patterns
• Use information-theoretical techniques to build behavior models of end hosts and applications
• Apply dominant state analysis to explain traffic behavior
• Discover typical behavior profiles as well as rare and deviant behaviors

References