The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

Automatic Detection of Emerging Threats to Computer Networks

Andre McDonald

CSIR
our future through science
CELEBRATING 70 Years
Ideas that work

Overview of presentation

- The cybercrime problem
  - Background
  - Cybercrime in South Africa
  - Network vulnerabilities
- Network intrusion detection
  - The need for network intrusion detection
  - Network intrusion detection systems
  - Anomaly detection techniques
- CSIR research and development
  - Time series detection
  - Network intrusion detection software platform

The Cybercrime Problem

The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

CSIR
our future through science
CELEBRATING 70 Years
Ideas that work

Background

Cybercrime is any crime in which a computer is the object of the crime, or is used as a tool to commit an offence.

Examples

- Fraud and financial crimes
- Identity theft and theft of classified information

With a global impact of $388 bn per year, cybercrime is

- bigger than the global black market in marijuana, cocaine and heroin combined ($288 bn), and
- more than 100 times the annual expenditure of UNICEF ($3.7 bn).

Cybercrime in South Africa

2014 statistics for cybercrime in South Africa	
Estimated cost to SA companies	R 5.8 billion, 0.14% GDP
Average delay in detecting breach	200 days
Online adult South Africans exposed to cybercrime	55%
Global ranking in cybercrime exposure	3rd, after Russia and China
Estimated rate of “phishing” attacks	1 in 215 email messages

Challenges in the local context

- SMMEs with small ICT budgets
- Local skills shortage and lack of awareness

Network vulnerabilities

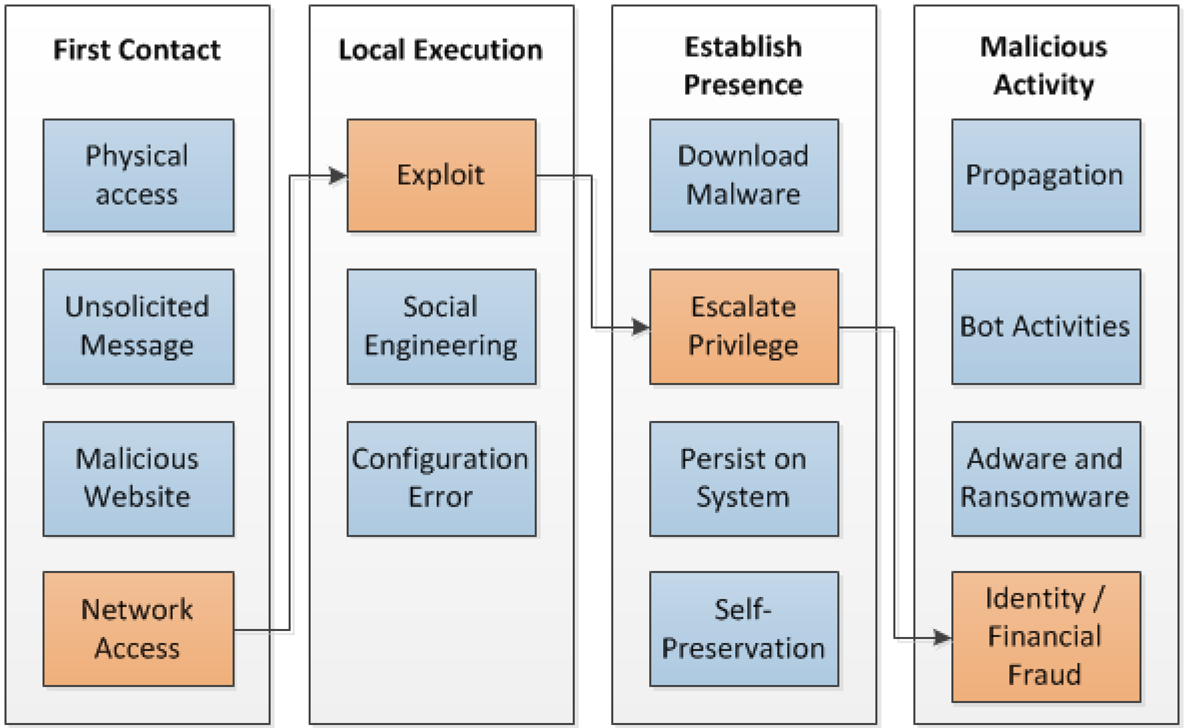
- Cybercrime frequently involves illegitimate access to networked computer systems, or their abuse
- Networked systems are vulnerable
  - Weak passwords
  - Mobile devices and BYOD policies
  - Outdated software and misconfiguration
  - Unencrypted transmission of information
- Emerging threats and previously unobserved attacks are of particular concern
  - Rapid threat propagation and slow reaction
  - Threat signatures unavailable or not up to date

Network Intrusion Detection

The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

The need for network intrusion detection

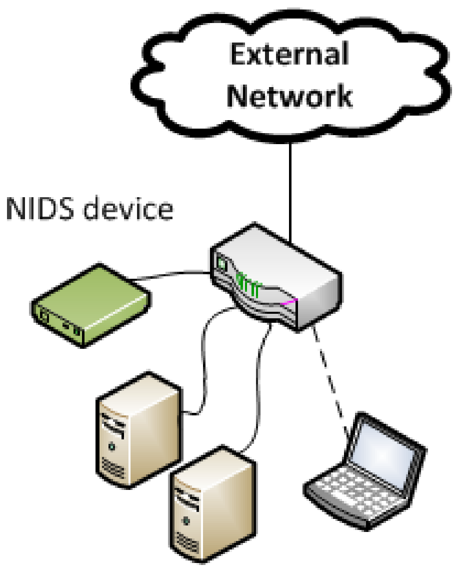
- **Automatic detection of network intrusions is required as an additional security layer**
- Timely detection and blocking of intrusions in their early phases may limit the scope of the damage

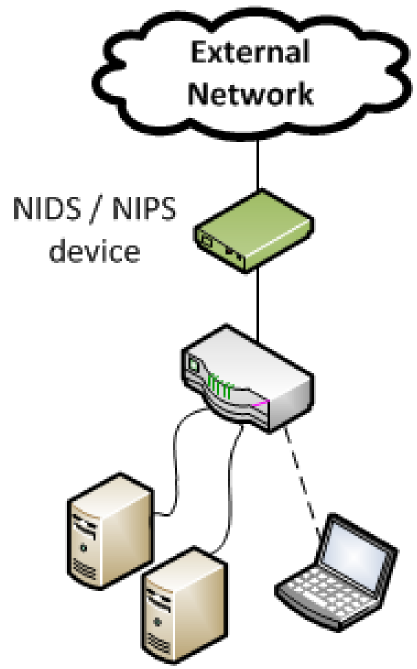


Network intrusion detection systems

Network Intrusion Detection Systems (NIDS):

Hardware or software systems for automatically detecting intrusions in computer networks





Detection approaches

Misuse detection

- Select predefined signatures of known malicious traffic patterns
- Compare observed traffic to signatures
- Low false positive rate, but cannot detect previously unobserved attacks

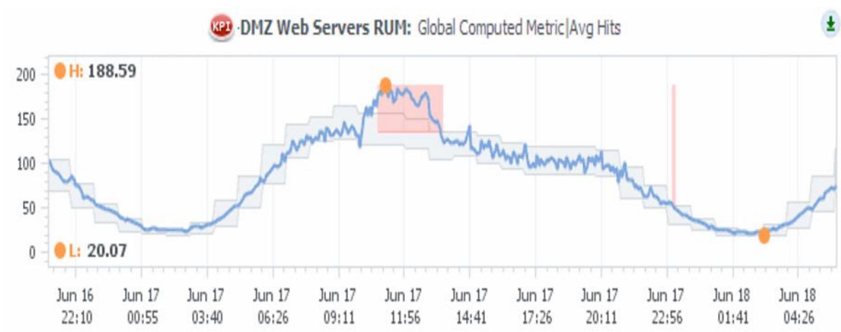
Anomaly detection

- Construct models of legitimate traffic patterns
- Observed traffic that deviates from models are tagged as malicious
- Can detect certain previously unobserved attacks, but typically exhibits high false positive rates

Anomaly detection techniques

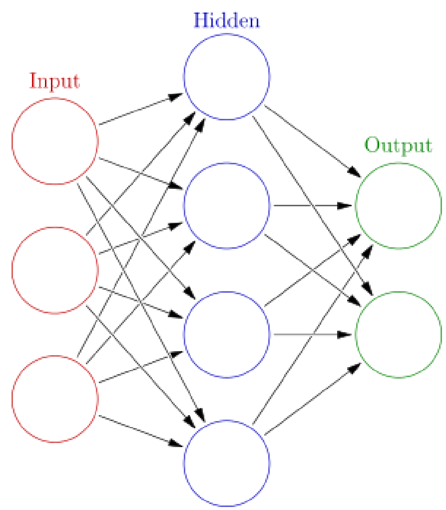
Statistical techniques

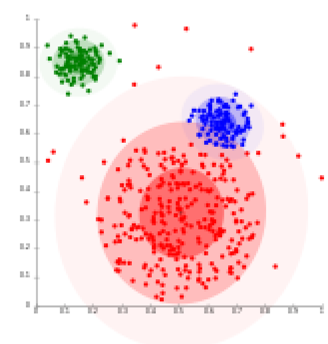
- Univariate / multivariate models
- Time series detection
  - Filtering and thresholding



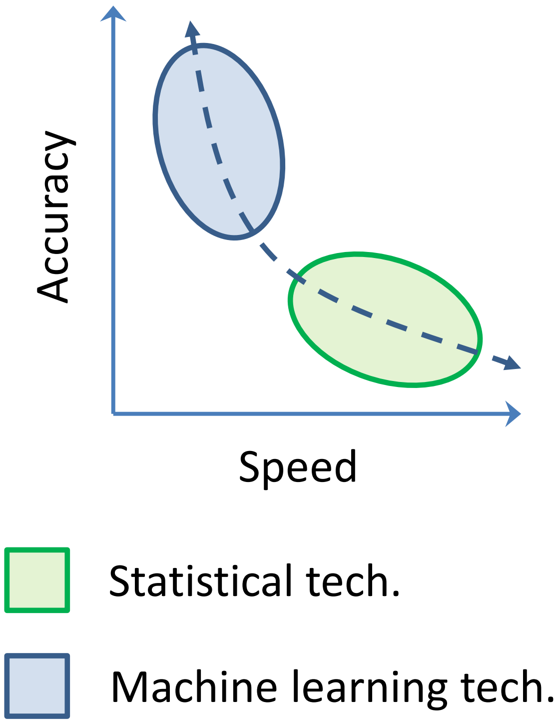
Machine learning techniques

- Supervised learning
  - SVMs, neural networks
- Unsupervised learning
  - Clustering, outlier detection





Anomaly detection techniques



Statistical techniques

- Adaptive and online model construction
- Can potentially be trained by attacker
- Rapid detection implies more false positives

Machine learning techniques

- Ability to detect more intricate attacks
- Heavy data pre-processing burden
- Supervised learning requires labelled data

CSIR Research and Development

The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

CSIR
our future through science
CELEBRATING 70 Years
Ideas that work

CSIR anomaly detection research and development

Statistical time series based detectors

- Two-stage detection
- Multi-resolution detection

} Suppression of false positives

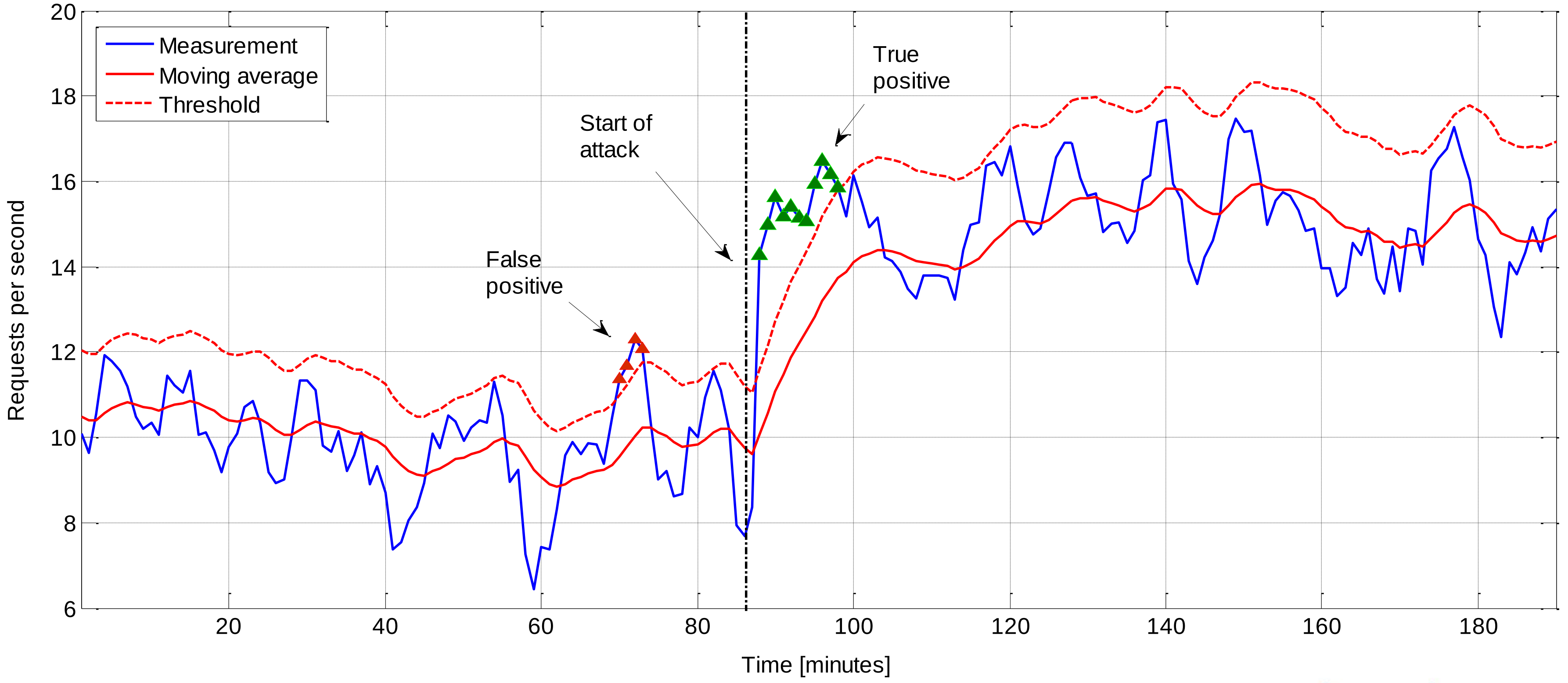
Unsupervised machine learning based detectors

- Unsupervised feature selection to address lack of labelled data

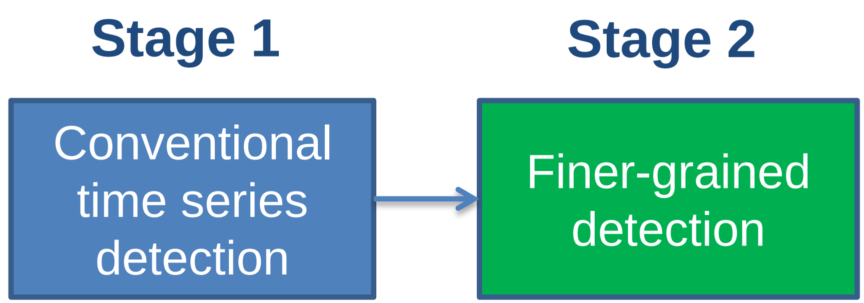
Network intrusion detection software platform development

- Deployment and configuration of sensors and detectors
- Monitoring of network traffic patterns
- Incident logging and analysis

Time series detection: Example

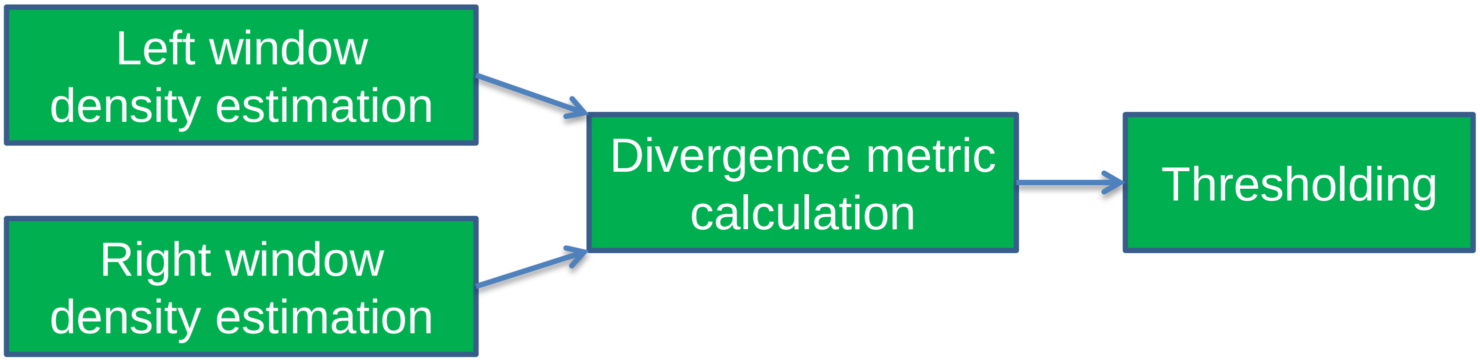


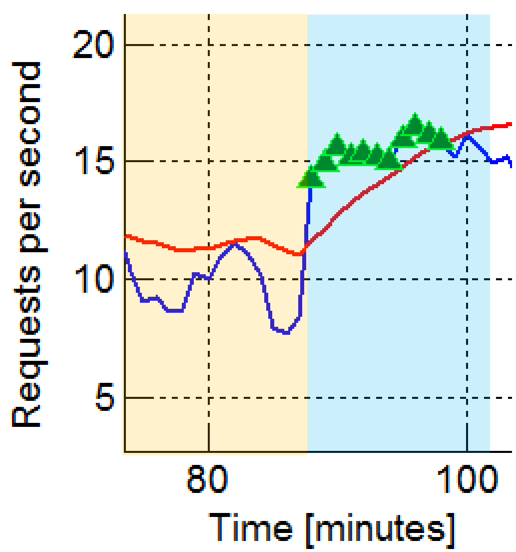
Time series detection: Proposed two-stage detector

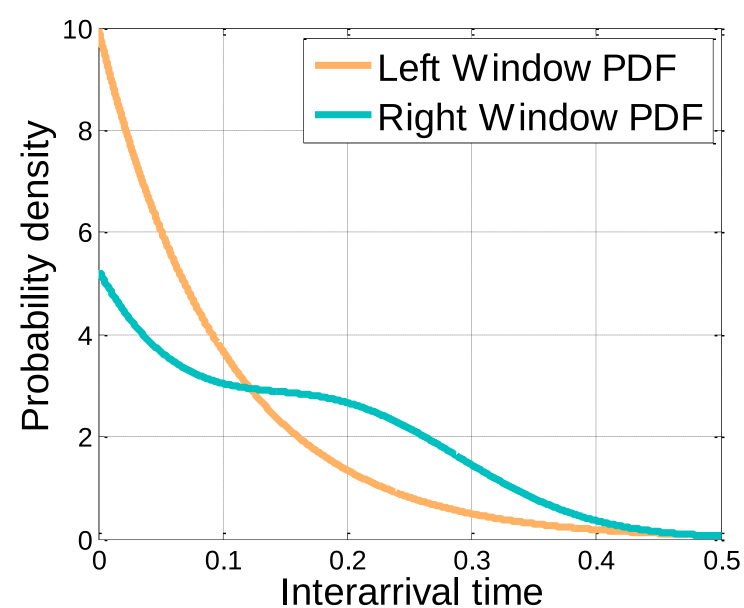


- Second stage performs finer-grained detection to suppress false positives
- Second stage algorithm triggers only upon threshold crossing in first stage detector
- Candidate algorithms for stage 2:
  - Spectral-based detector
  - **Inter-arrival time detector**

Time series detection: Proposed two-stage detector



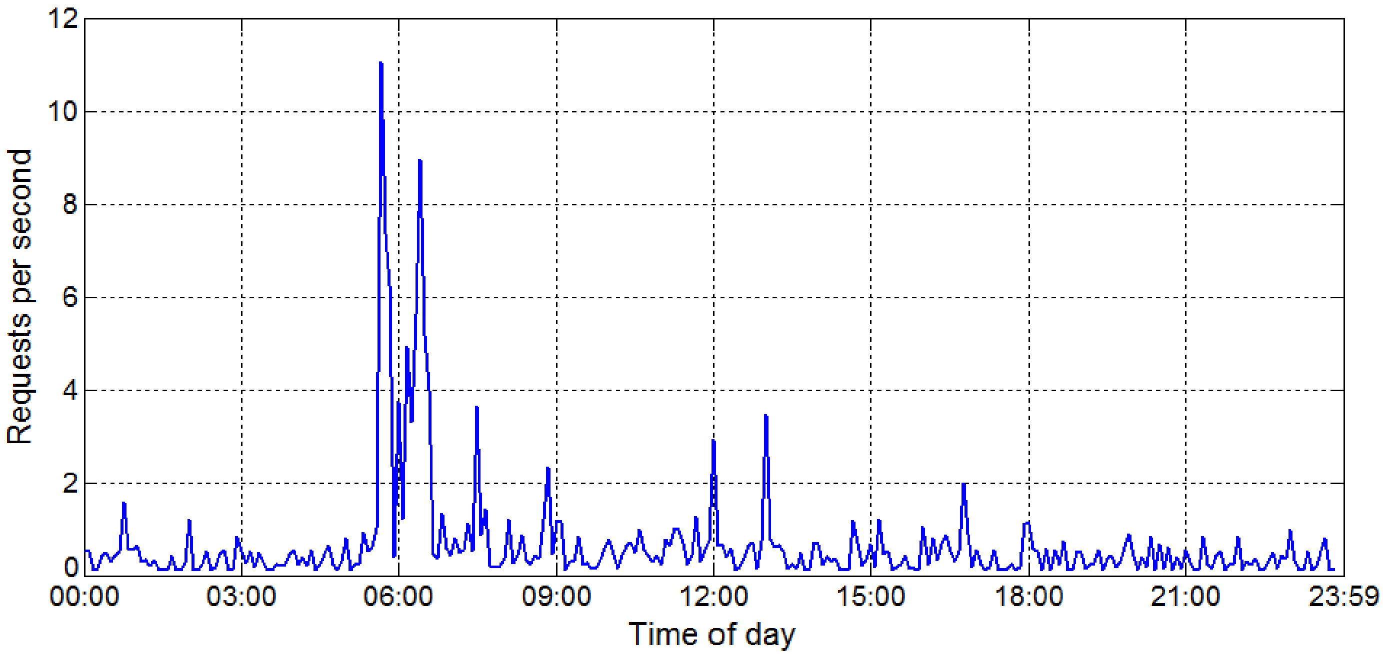




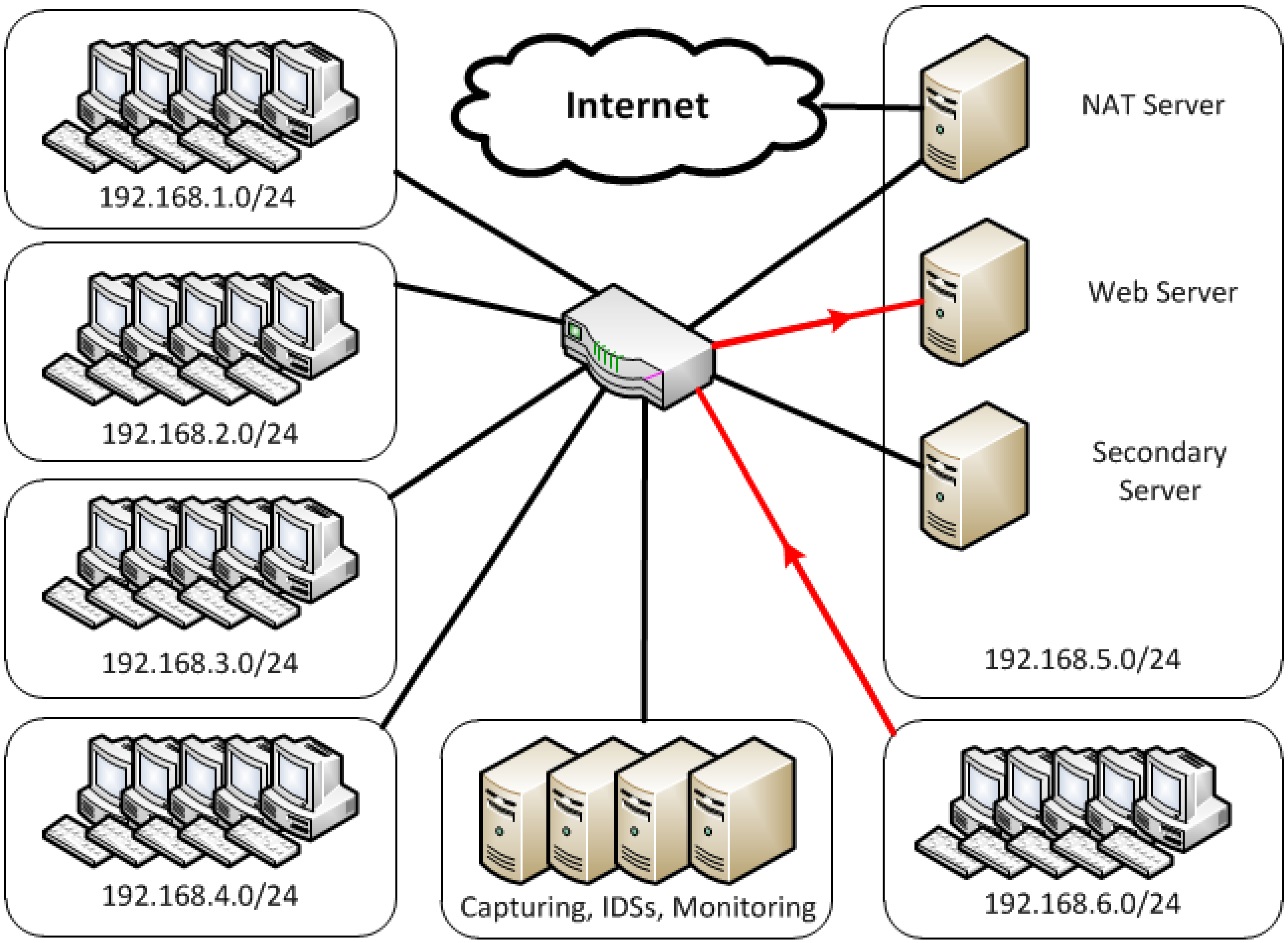
Time series detection: Experimental work

Detection of denial-of-service attack against web server

- Corporate network with compromised workstations
- Compromised workstation floods web server with requests, denying legitimate users access to the website



Time series detection: Experimental work



Time series detection: Experimental work

The 5th CSIR CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

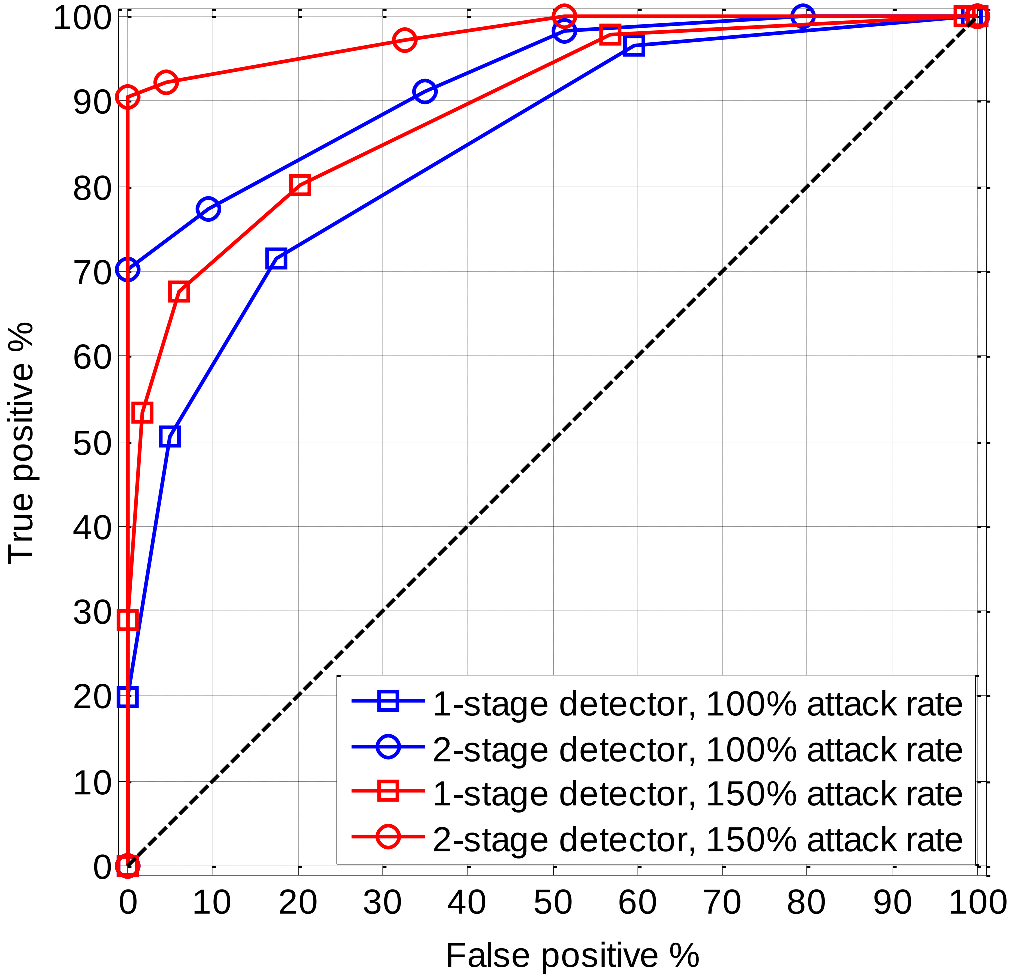
Stage	Algorithm / Parameter	Value
1	Detection	Exponentially-weighted moving average
1	Bin width	2 seconds
2	Density estimation	Gaussian kernel, Silverman's heuristic over logarithm of request interarrival time
2	Divergence metric	Symmetric Kullback-Leibler divergence
2	Window width	31 requests

$$J(\hat{f}_L(t), \hat{f}_R(t)) \triangleq D_{KL}(\hat{f}_L || \hat{f}_R) + D_{KL}(\hat{f}_R || \hat{f}_L)$$

$$D_{KL}(\hat{f}_L || \hat{f}_R) \triangleq \int_{-\infty}^{\infty} \hat{f}_L(t) \ln \left[\frac{\hat{f}_L(t)}{\hat{f}_R(t)} \right] dt$$

Time series detection: Experimental results

The 5th CSIR CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC



Thank you