Covert Channel Detection Using Process Query Systems

➡ **Annarita Giani**
Vincent Berk
George Cybenko

Institute for Security Technology Studies
Thayer School of Engineering
Dartmouth College
Hanover, NH

FLoCon 2005

MOTIVATION

CNN.COM
Sunday, June 19, 2005 Posted: 0238 GMT (1038 HKT)

NEW YORK -- The names, banks and account numbers of up to 40 million credit card holders may have been accessed by an unauthorized user, MasterCard International Inc. said.

Interest in network and computer security

Started investigating DATA EXFILTRATION

COVERT CHANNELS are the most subtle way of moving data.

They easily bypass current security tools.

Until now there has not been enough interest. So detection is still at the first stage.

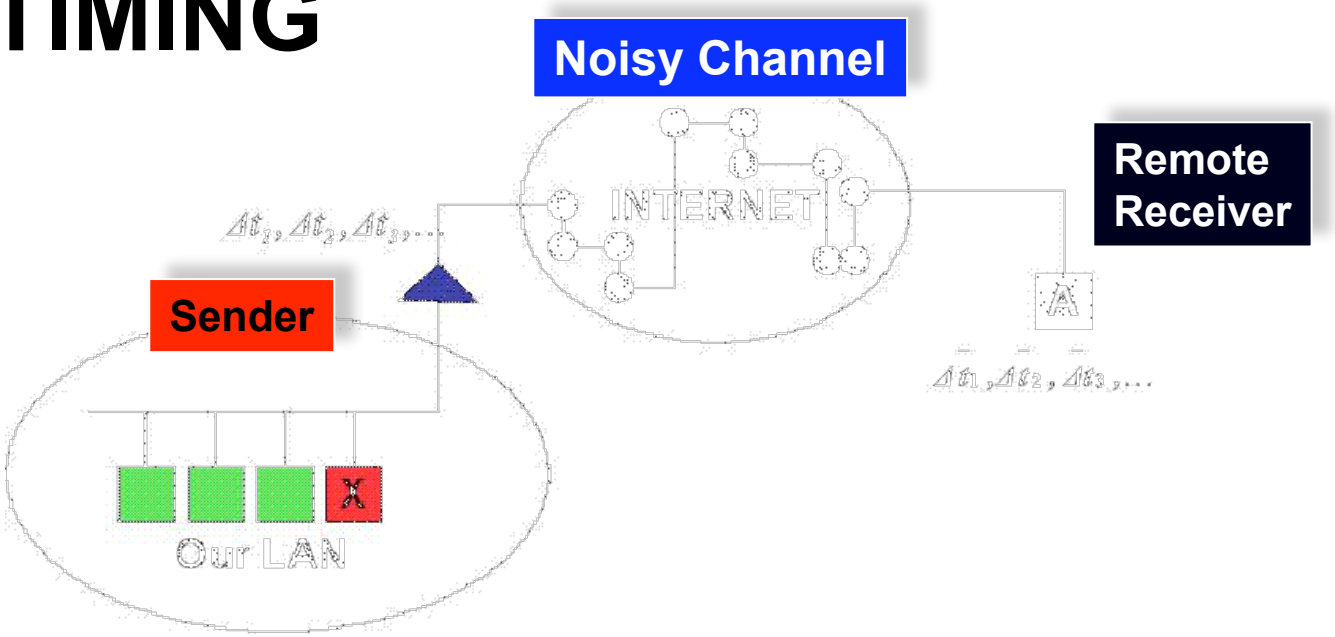
OUTLINE

- **Covert Channels**
- **Process Query Systems**
- **Detection of covert channels using a PQS**

“A communication channel is covert if it is neither designed nor intended to transfer information at all.” (Lampson 1973)

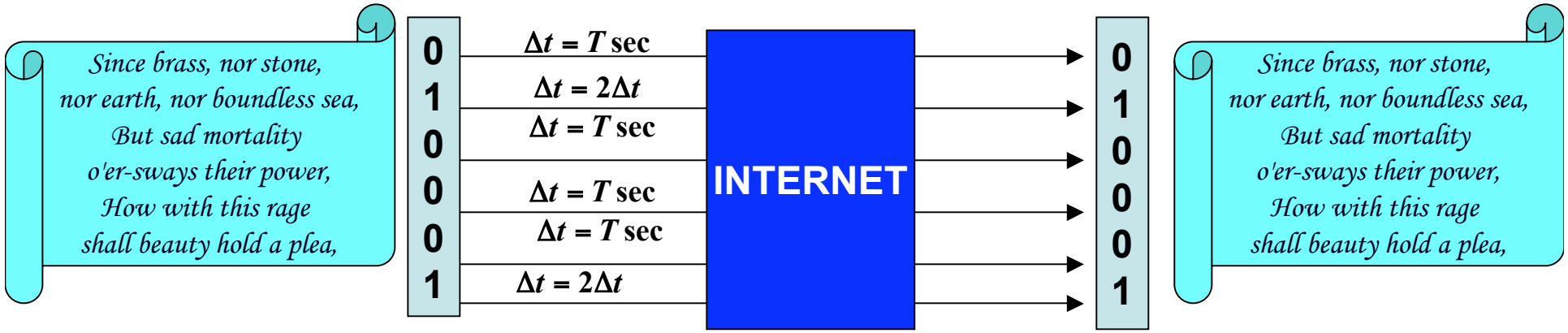
“Covert channels are those that use entities not normally viewed as data objects to transfer information from one subject to another.” (kemmerer 1983)

EXAMPLE: TIMING COVERT CHANNEL



Two approaches

1. **Information Theory**
2. **Statistical analysis**

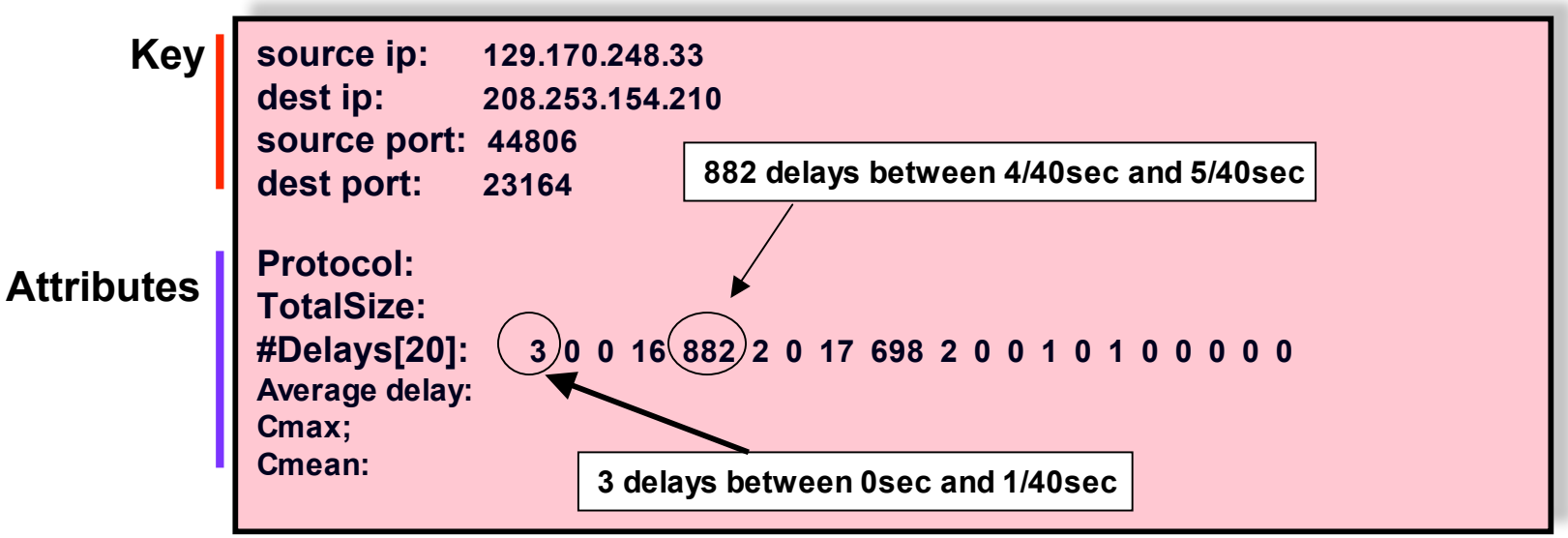


Sensor

Traffic is separated in connection types

We built a package that registers the time delays between consecutive packets for every network traffic flow.

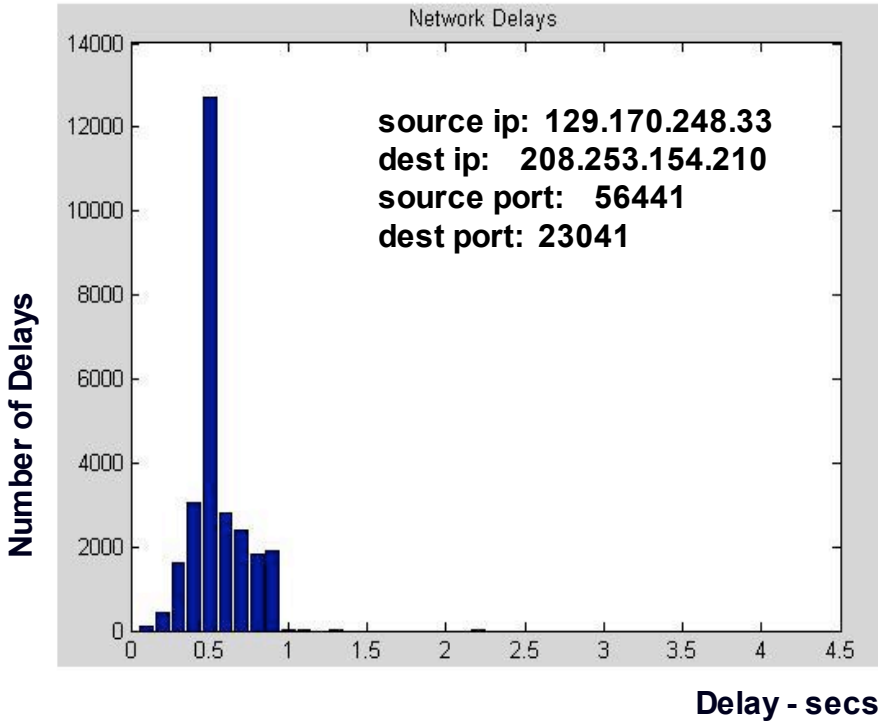
Given an interval of time we build the following node:

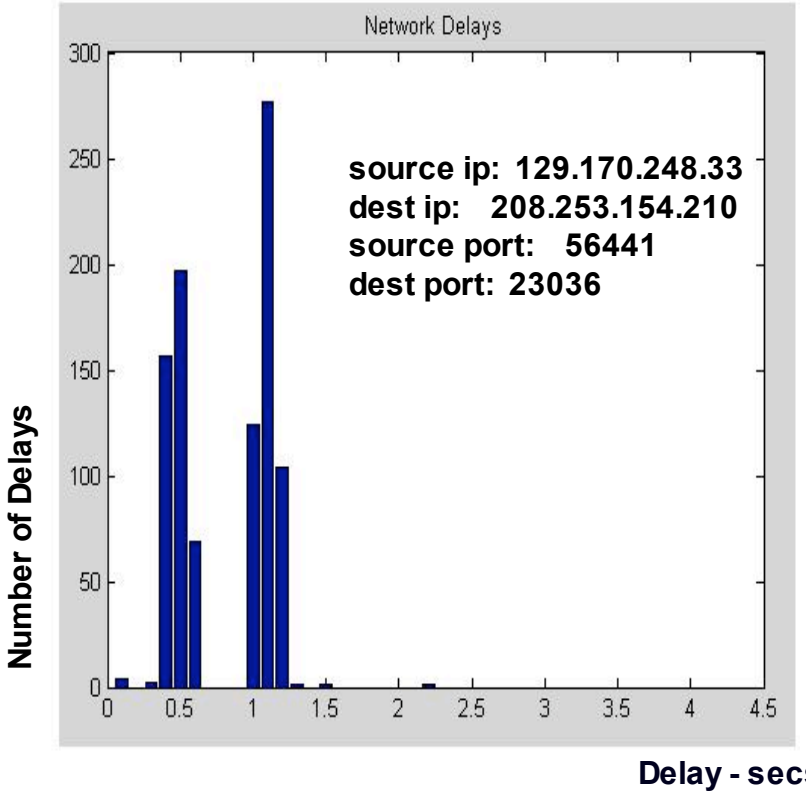


Covert Channels

Assumptions of the experiments:

- **No malicious noise.**
- **Binary source.**





OUTLINE

- Covert Channels
- ➡ Process Query Systems
- Detection of covert channels using a PQS

Process Query Systems for Homeland Security

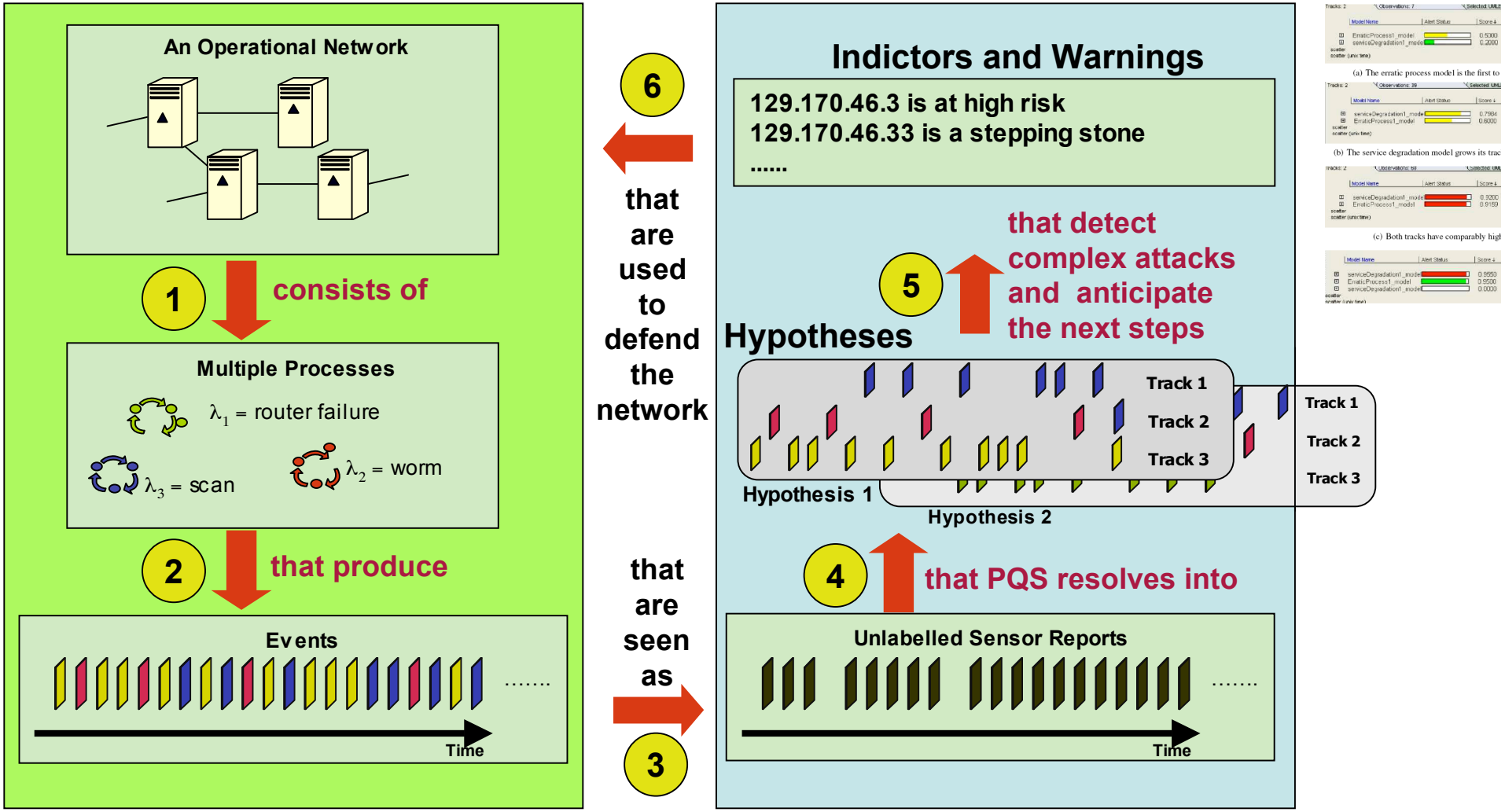
- **How it works:**
  - User provides a *process* description as query
  - PQS monitors a stream of sensor data
  - PQS matches sensor data with registered queries
  - A match indicates that the process model may explain that sensor data, hence that process may be the cause of those sensor readings.

Applications

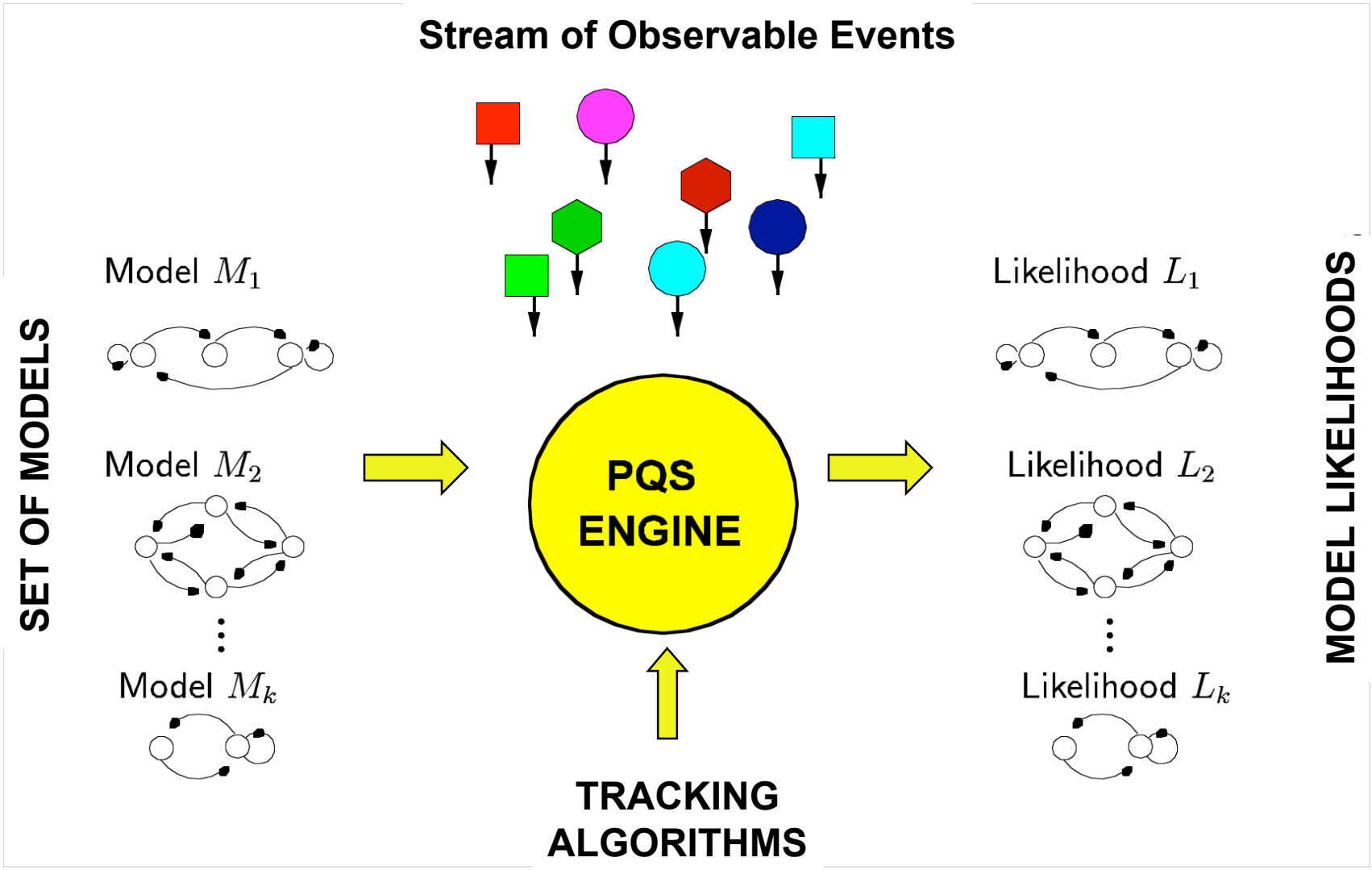
- **Tactical C4ISR** - Is there a large ground **vehicle** convoy moving towards our position?
- **Cyber-security** - Is there an unusual pattern of **network and system calls** on a server?
- **Autonomic computing** - Is my **software** operating normally?
- **Plume detection** – where is the source of a hazardous **chemical plume**?
- **FishNet** – how do **fish** move**?**
- **Insider Threat Detection** - Is there a pattern of unusual **document accesses** within the enterprise document control system?
- **Homeland Security -** Is there a pattern of unusual **transactions**?
- **Business Process Engineering** - Is the **workflow system** working normally?
- **Stock Market**
- **…**

All are "adversarial" processes, not cooperative so the observations are not necessarily labeled for easy identification and association with a process!

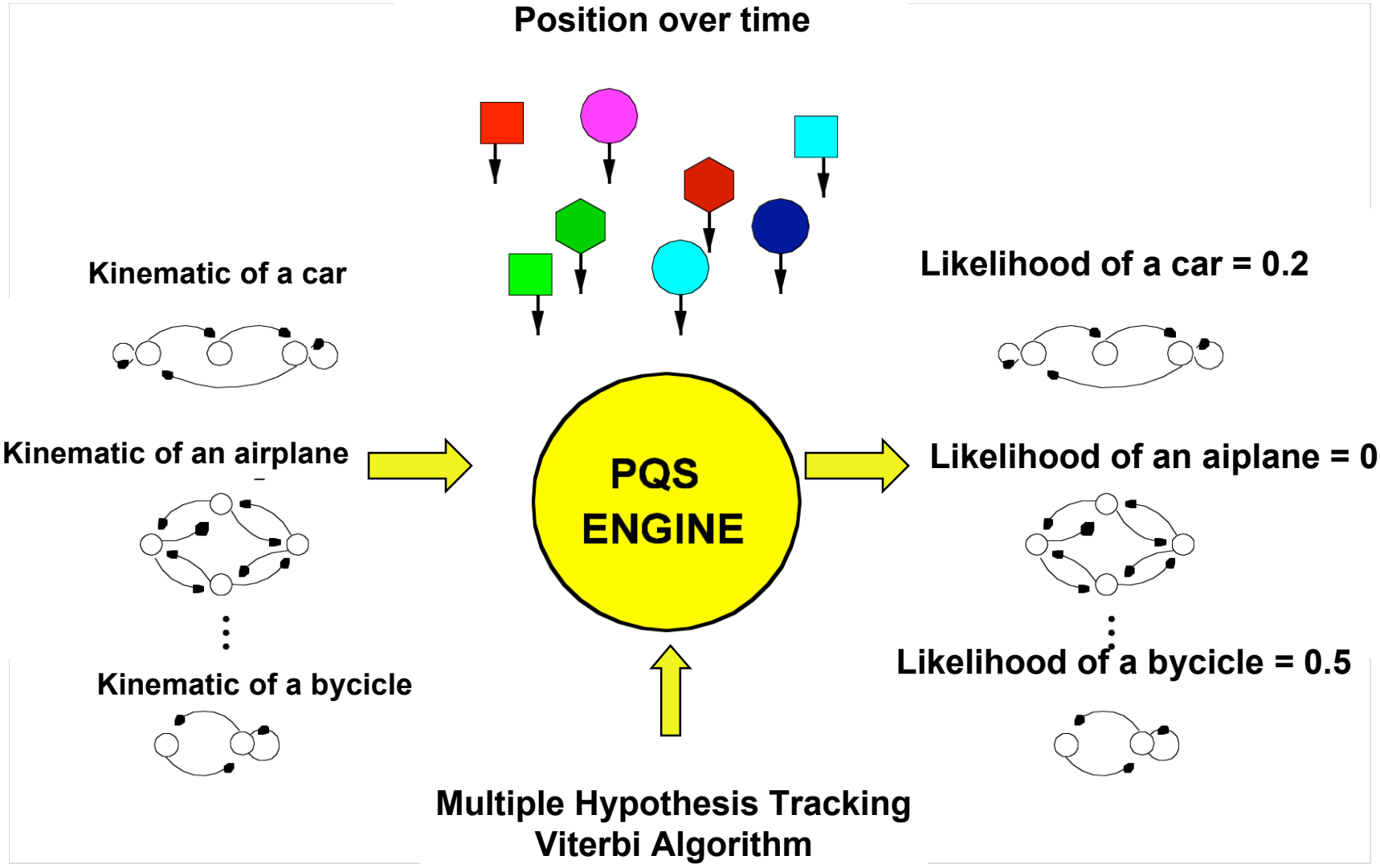
Example


An Operational Network
1
consists of
Multiple Processes
λ1 = router failure
λ3 = scan
λ2 = worm
2
that produce
Events
.....
Time
6
that are used to defend the network
that are seen as
3
Indictors and Warnings
129.170.46.3 is at high risk
129.170.46.33 is a stepping stone
.....
5
that detect complex attacks and anticipate the next steps
Hypotheses
Track 1
Track 2
Track 3
Hypothesis 1
Hypothesis 2
4
that PQS resolves into
Unlabelled Sensor Reports
.....
Time

PQS


Stream of Observable Events
SET OF MODELS
Model M_1
Model M_2
Model M_k
PQS ENGINE
TRACKING ALGORITHMS
Likelihood L_1
Likelihood L_2
Likelihood L_k
MODEL LIKELIHOODS

PQS


Position over time
Kinematic of a car
Kinematic of an airplane
Kinematic of a bycicle
PQS ENGINE
Likelihood of a car = 0.2
Likelihood of an aiplane = 0.01
Likelihood of a bycicle = 0.5
Multiple Hypothesis Tracking
Viterbi Algorithm

OUTLINE

- Covert Channels
- Process Query Systems
- ➡ Detection of covert channels using a PQS

Observations

Time T

source ip: 129.170.248.33
dest ip: 208.253.154.210
source port: 44806
dest port: 23164

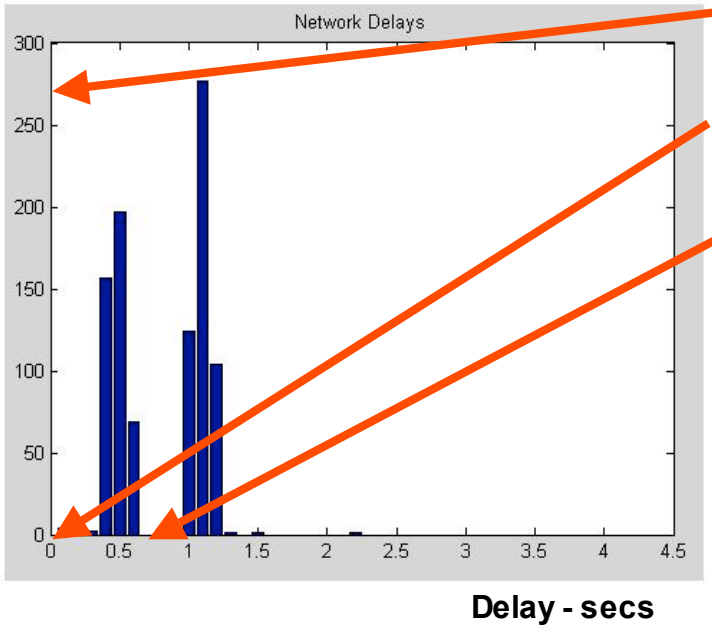
C_{mean}/C_{max}

Time T+1

source ip: 129.170.248.33
dest ip: 208.253.154.210
source port: 44806
dest port: 23164

C_{mean}/C_{max}

Covert Channels models

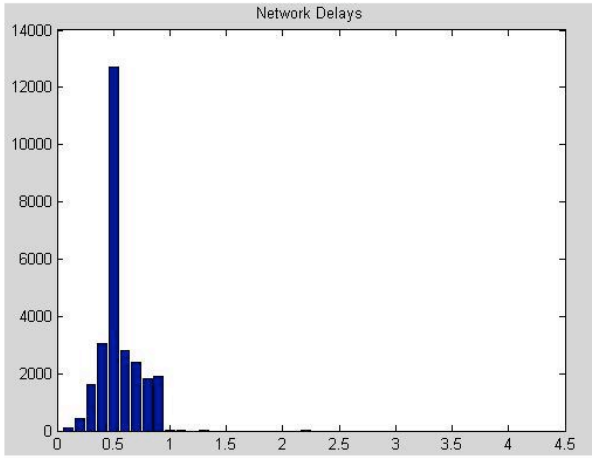


C_{max} = max number of packets with the same delay (= 280)

$C(i)$ = number of packets with interpacket delay (= 0)

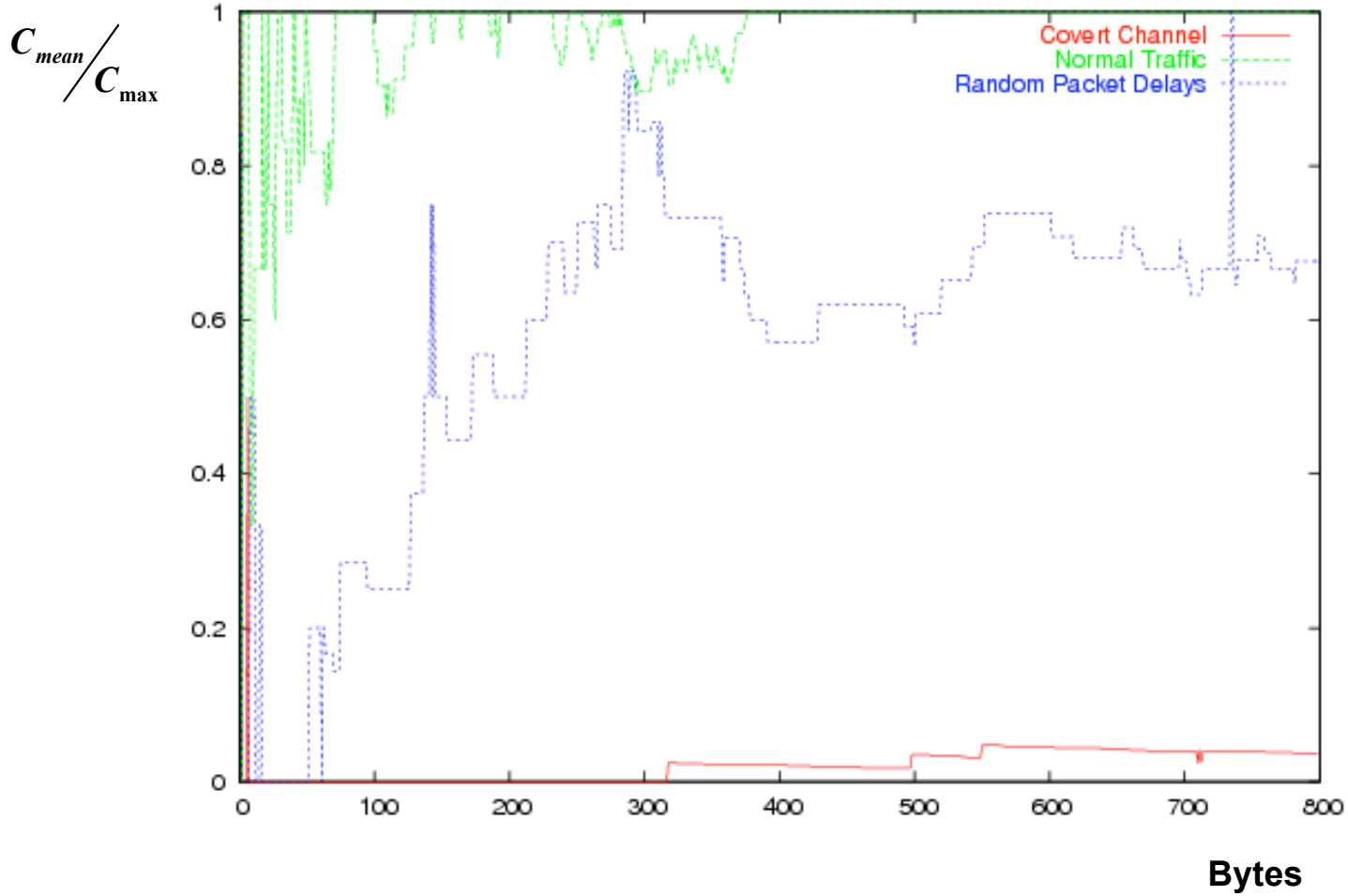
μ = sample mean of interpacket transmission times (= 0.7)

$$\frac{C(\mu)}{C_{max}} \approx 1 \Rightarrow$$ Not Covert Channel

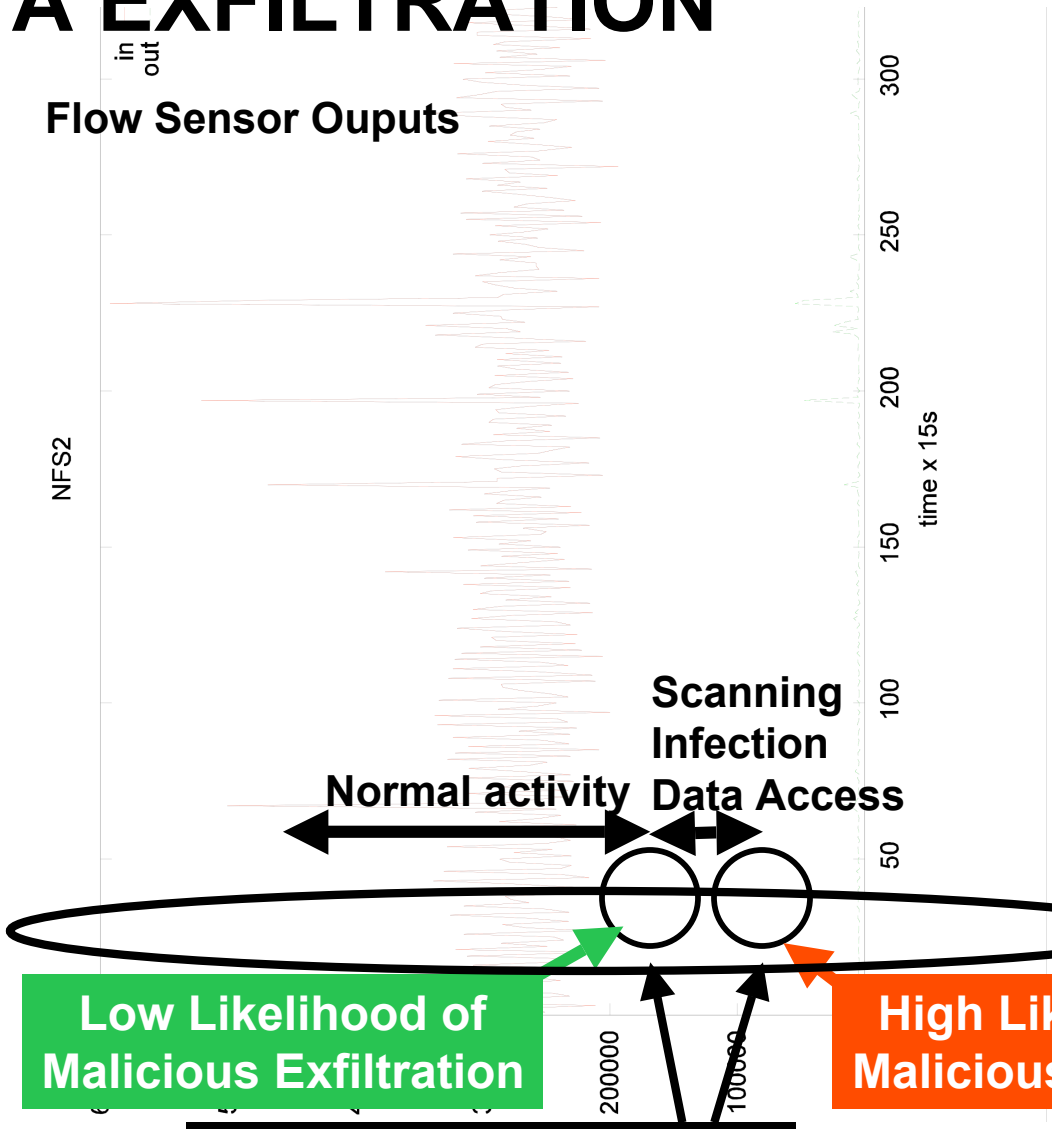


$$\frac{C(\mu)}{C_{max}} << 1 \Rightarrow$$ Covert Channel

RESULTS


C_{mean}/C_{max}
Covert Channel
Normal Traffic
Random Packet Delays
1
0.8
0.6
0.4
0.2
0
0
100
200
300
400
500
600
700
800
Bytes

DATA EXFILTRATION

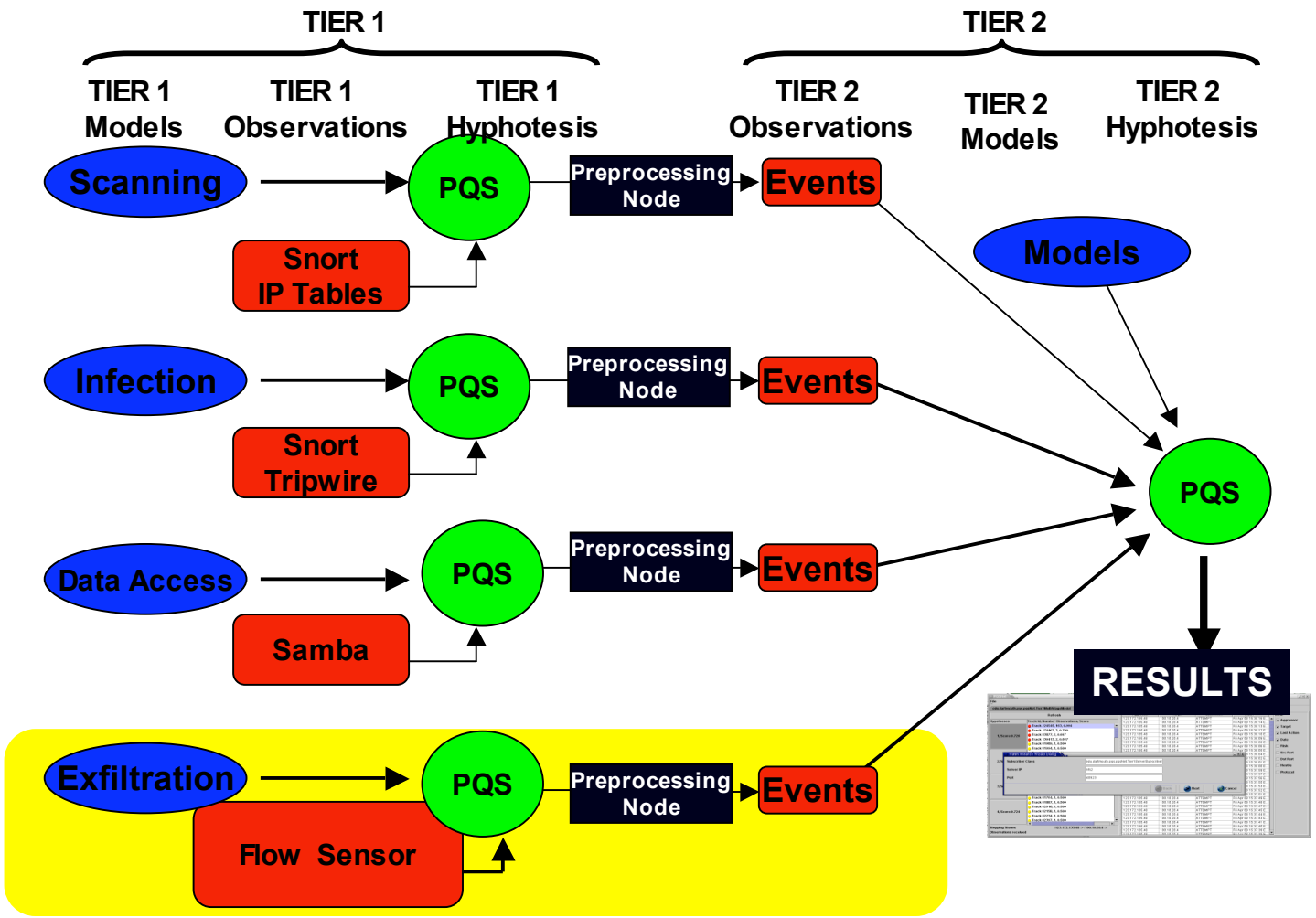


Exfiltration modes:

- SSH
- HTTP
- FTP
- Email
- Covert Channel
- Phishing
- Spyware
- Pharming
- Writing to media
  - paper
  - drives
- etc

Also monitor inter-packet delays for covert channels

Hierarchical PQS Architecture



For more information :

www.pqsnet.net
www.ists.dartmouth.edu

annarita.giani@dartmouth.edu

vincent.berk@dartmouth.edu

george.cybenko@dartmouth.edu

Thanks.