Detecting Deception in the Context of Web 2.0.

Annarita Giani,

EECS, *University of California, Berkeley, CA*

Paul Thompson,

CS Dept. *Dartmouth College, Hanover, NH*

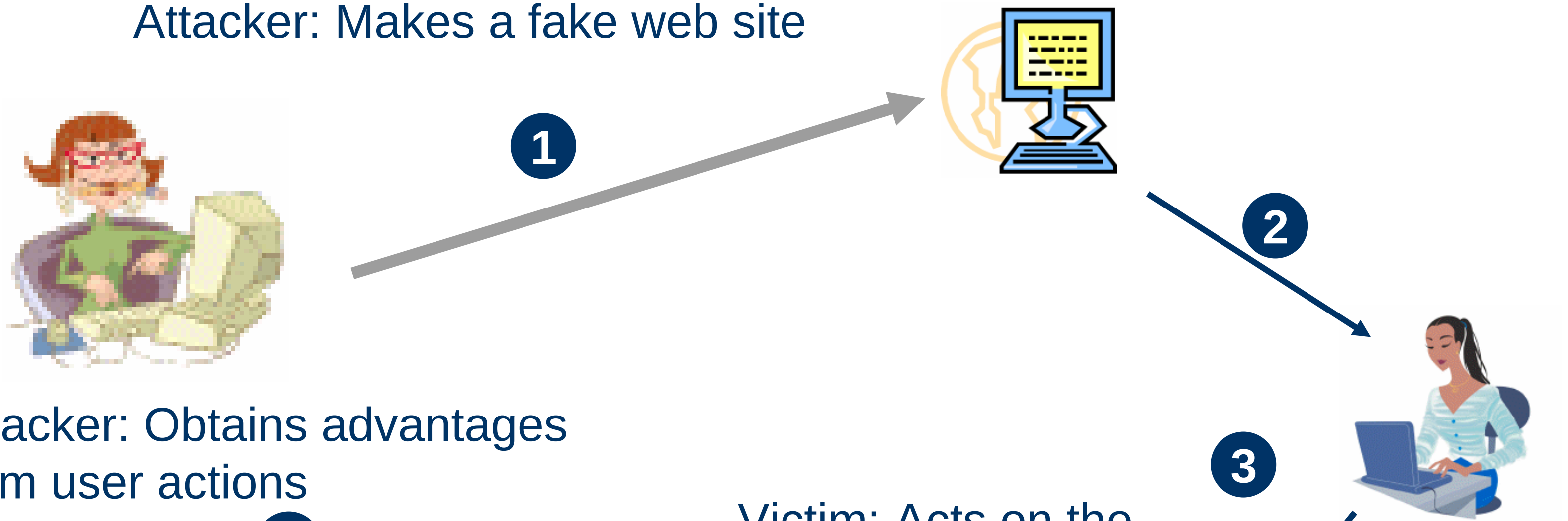
W2SP2007 – Oakland, CA – May 24, 2007

Outline

1. Motivation and Terminology
2. Process Query System (PQS) Approach
3. Detection of a complex attack
4. Conclusion and Acknowledgments

Cognitive Hacking

The user's attention is focused on the channel. The attacker exploits this fact and uses malicious information in the channel to mislead her.



Misleading information from a web site

Attacker: Makes a fake web site

Attacker: Obtains advantages from user actions

Victim: Acts on the information from the web site

MISINFORMATION - Lebed case

Jonathan Lebed.

He spread fake rumors about stocks.

Investors driven to buy shares of
that stock inflating its price

The SEC wanted to prosecuted him for stock fraud.

Was allowed to keep $500,000 from his "illegal" stock proceeds.

The law ???

"Subj: THE MOST UNDERVALUED STOCK EVER
"Date: 2/03/00 3:43pm Pacific Standard Time
"From: LebedTG1

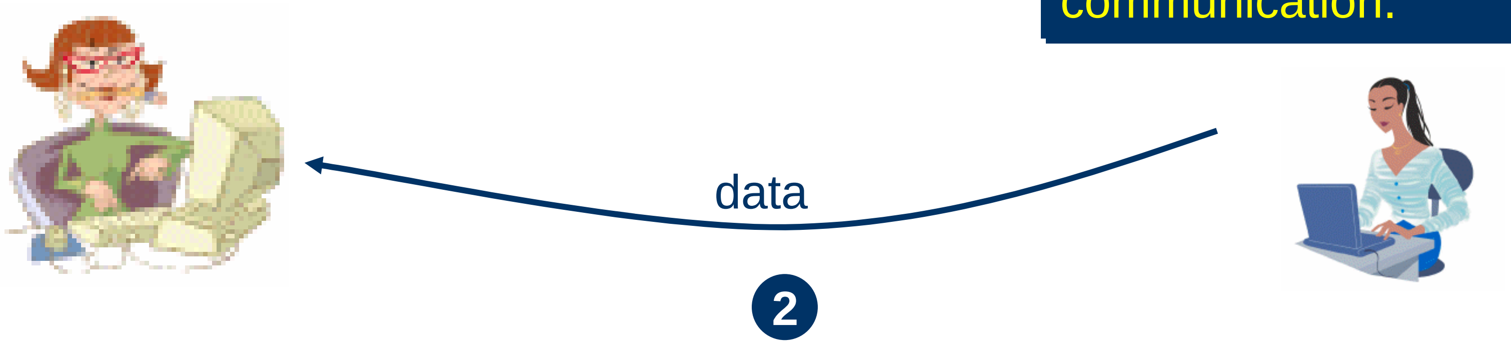
> "FTEC is starting to break out! Next week, this thing will EXPLODE. . . .
> "Currently FTEC is trading for just $2 1/2! I am expecting to see FTEC at $20 VERY SOON.
> "Let me explain why. . . .
> "The FTEC offices are extremely busy. . . . I am hearing that a number of HUGE deals are being worked on. Once we get some news from FTEC and the word gets out about the company . . . it will take-off to MUCH HIGHER LEVELS!
> "I see little risk when purchasing FTEC at these DIRT-CHEAP PRICES. FTEC is making TREMENDOUS PROFITS and is trading UNDER BOOK VALUE!!!"

Covert Channels

The user's attention is unaware of the channel. The attacker uses a medium not perceived as a communication channel to transfer information.

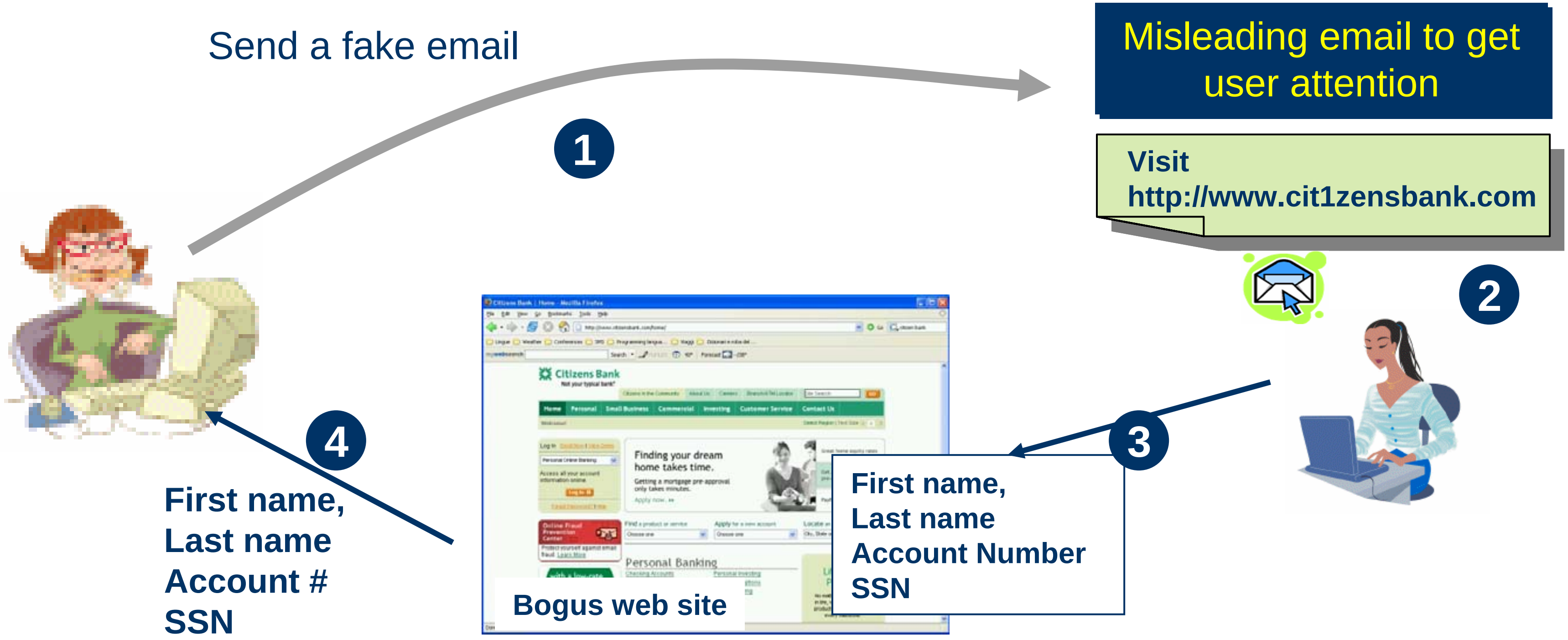
1 Attacker: Codes data into inter-packet delays, taking care to avoid drawing the attention of the user.

User: does not see inter-packet delay as a communication channel and does not notice any communication.



Phishing

The user's attention is attracted by the exploit. The information is used to lure the victim into using a new channel and then to create a false perception of reality with the goal of exploiting the user’s behavior.



Send a fake email

Misleading email to get user attention

Visit
http://www.cit1zensbank.com

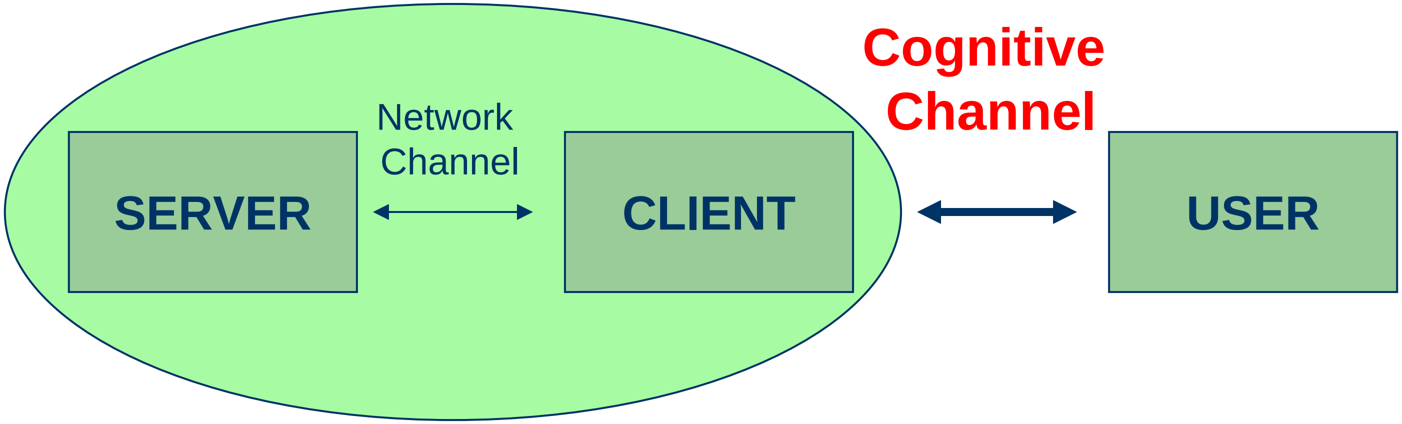
First name,
Last name
Account Number
SSN

Bogus web site

First name,
Last name
Account #
SSN

Cognitive Channels

A cognitive channel is a communication channel between the user and the technology being used. It conveys what the user sees, reads, hears, types, etc.



Focus of the current protection and detection approaches

The cognitive channel is the weakest link in the whole framework. Little investigation has been done on detecting attacks on this channel.

Cognitive Attacks

Our definition is from an engineering point of view.

Cognitive attacks are computer attacks over a cognitive channel. They exploit the attention of the user to manipulate her perception of reality and/or gain advantages.

COGNITIVE HACKING. The user's attention is focused on the channel. The attacker exploits this fact and uses malicious information to mislead her.

COVERT CHANNELS. The user is unaware of the channel. The attacker uses a medium not perceived as a communication channel to transfer information.

PHISHING. The user's attention is attracted by the exploit. The information is used to lure the victim into using a new channel and then to create a false perception of reality with the goal of exploiting the user's behavior.

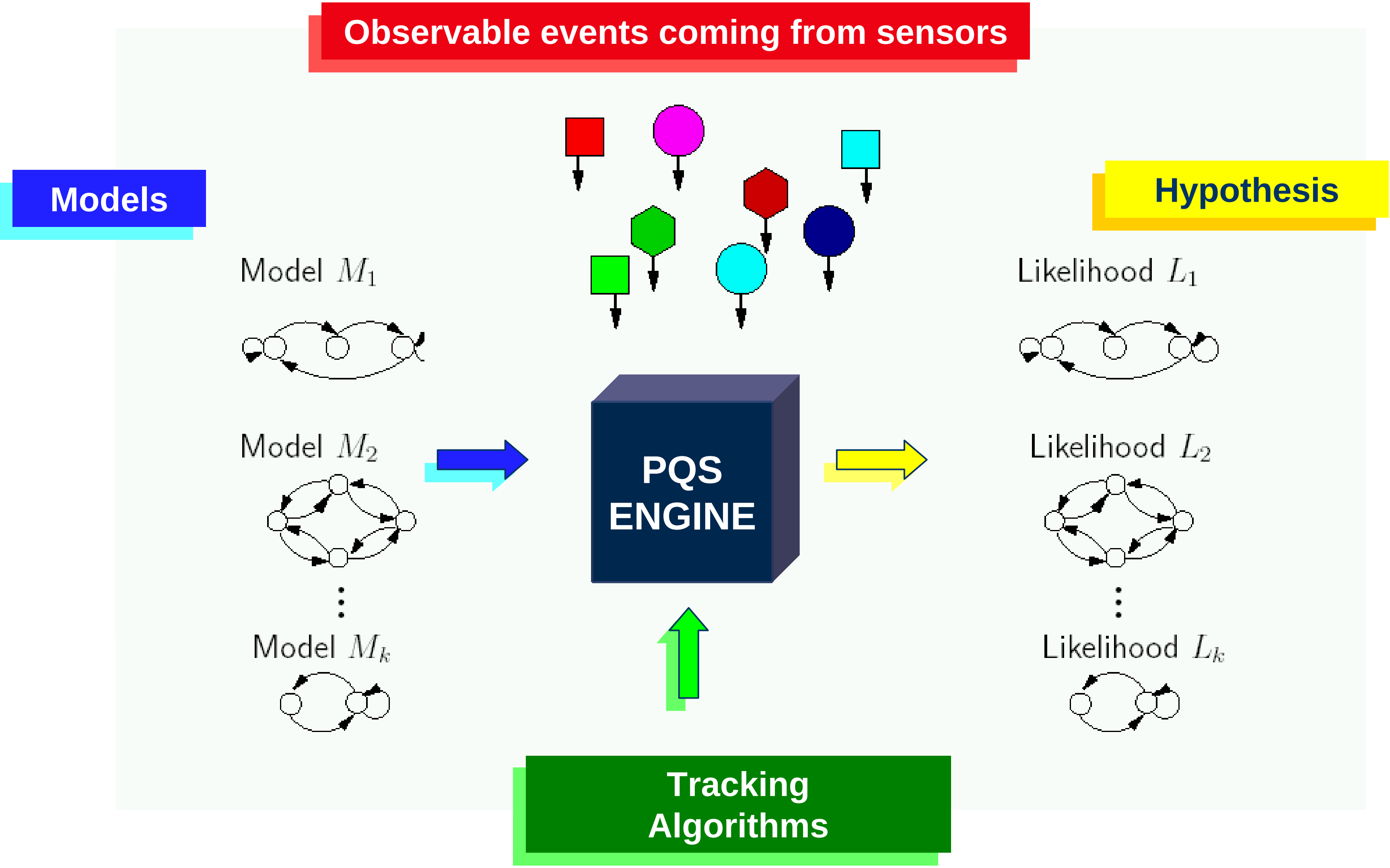
The Need to Correlate Events

- Large amount of sensors for network monitoring
  - Intrusion Detection Systems
  - Network traces
  - File Integrity Checkers
- Large amount of Alerts
  - Overloaded operators
  - Hard to make sense of alarms
- Need a principled way of combining alerts
  - Reduce false alarms
  - Discover multistage attacks

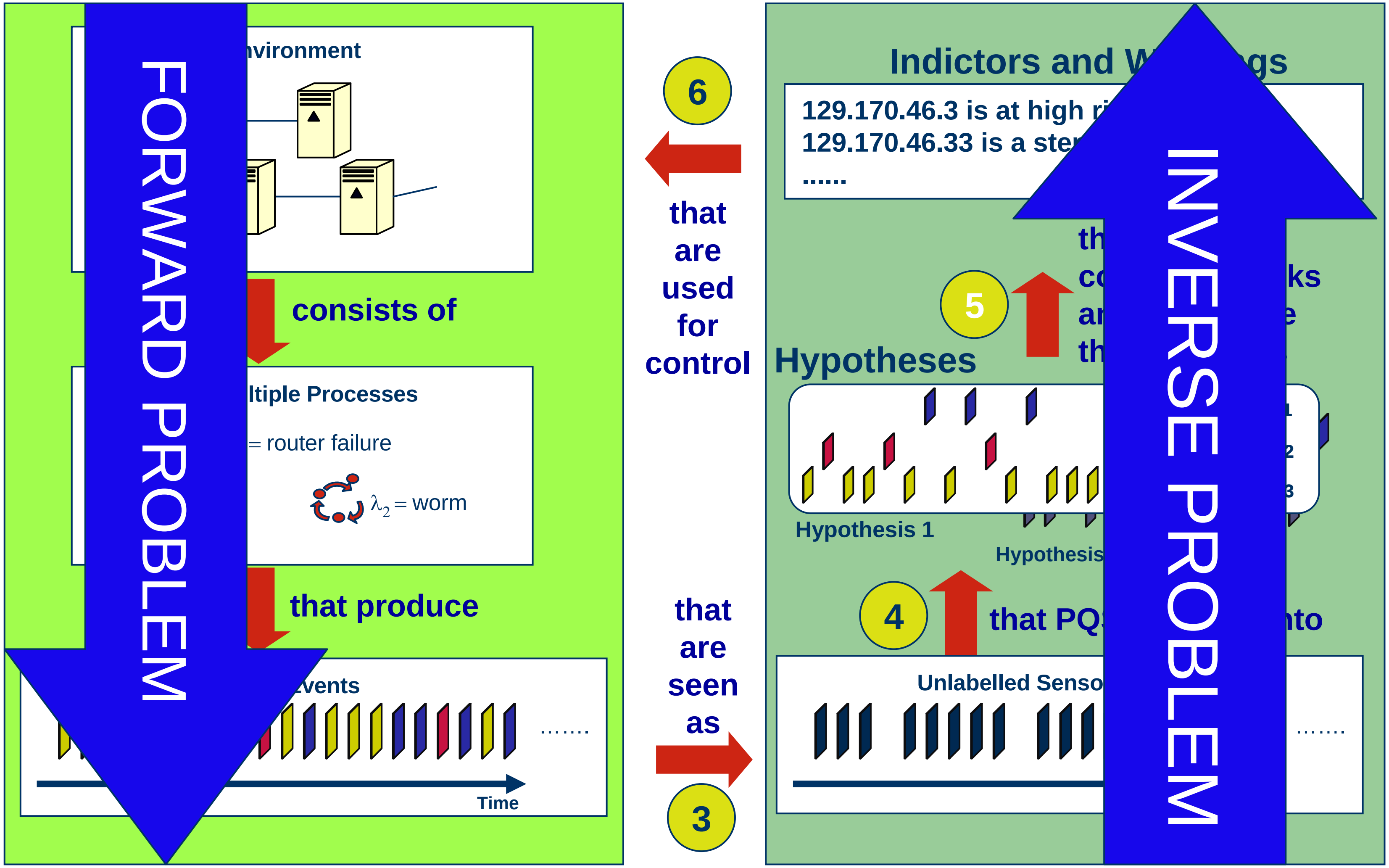
Outline

1. Motivation and Terminology
2. Process Query System (PQS) Approach
3. Detection of a complex attack
4. Conclusion and Acknowledgments

Process Query System



Framework for Process Detection



FORWARD PROBLEM

nvironment

consists of

ltiple Processes

= router failure

λ_2 = worm

that produce

vents

....

Time

Real World

6

that are used for control

that are seen as

3

INVERSE PROBLEM

Indictors and W gs

129.170.46.3 is at high ri

129.170.46.33 is a ste

......

th

co ks

ar e

th

5

Hypotheses

Hypothesis 1

Hypothesis

1

2

3

4

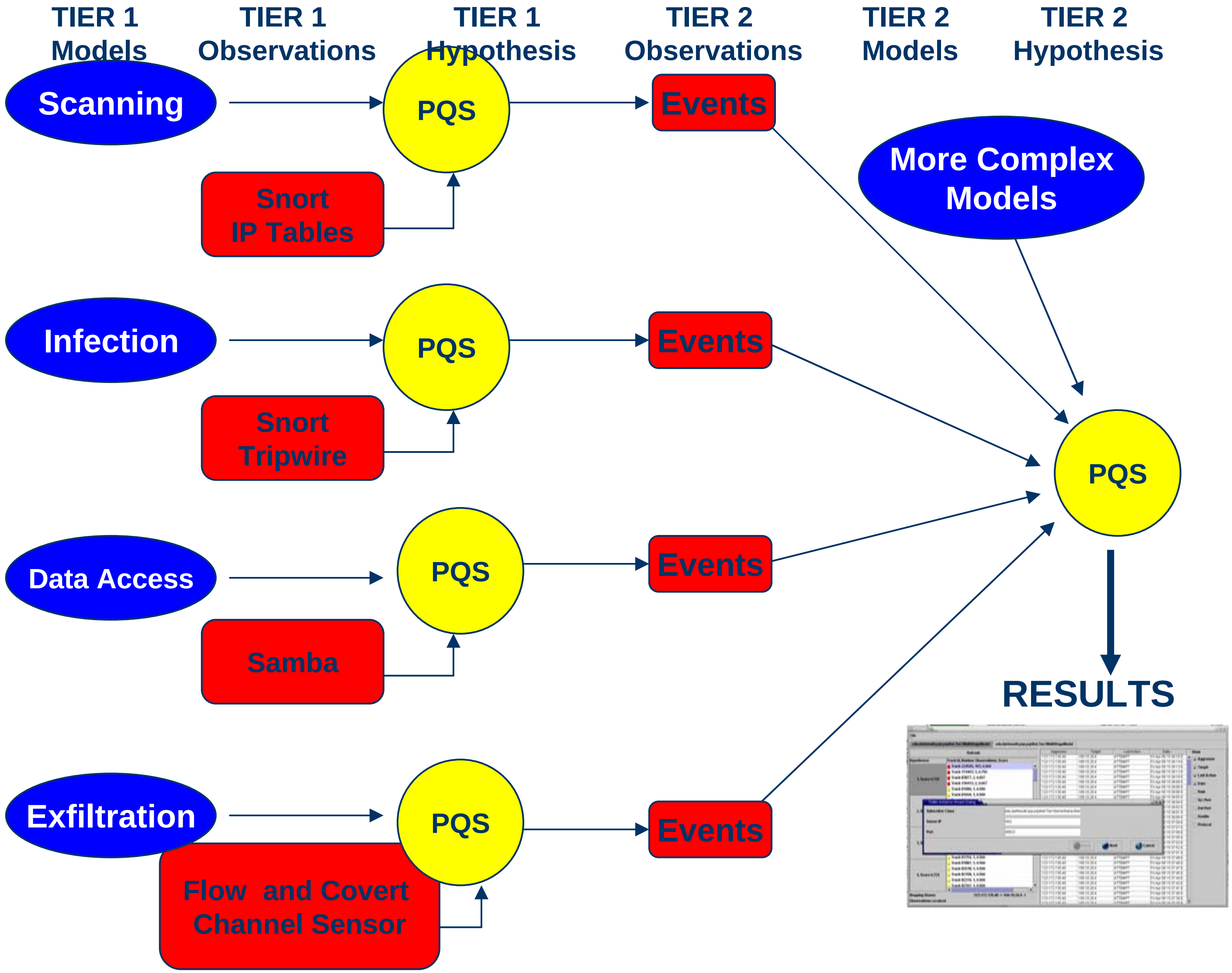
that PQ nto

Unlabelled Senso

....

Process Detection (PQS)

Hierarchical PQS Architecture



TIER 1 Models

TIER 1 Observations

TIER 1 Hypothesis

TIER 2 Observations

TIER 2 Models

TIER 2 Hypothesis

Scanning

PQS

Events

Snort
IP Tables

More Complex Models

Infection

PQS

Events

Snort
Tripwire

PQS

Data Access

PQS

Events

Samba

RESULTS

Exfiltration

PQS

Events

Flow and Covert Channel Sensor

Hidden Discrete Event System Models

Dynamical systems with discrete state spaces that are:

Causal - next state depends only on the past
Hidden – states are not directly observed
Observable - observations conditioned on hidden state are independent of previous states

Example. Hidden Markov Model

N States
M Observation symbols
State transition Probability Matrix, A
Observation Symbols Distribution, B
Initial State Distribution π

HDESM models are general

HDESM Process Detection Problem

Identifying and tracking several (casual discrete state) stochastic processes (HDESM's) that are only partially observable.

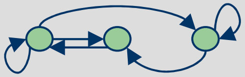
TWO MAIN CLASSES OF PROBLEMS

Hidden State Estimation: Determine the "best" hidden states sequence of a particular process that accounts for a given sequence of observations.

Discrete Sources Separation: :Determine the "most likely" process-to-observation association

Discrete Source Separation Problem

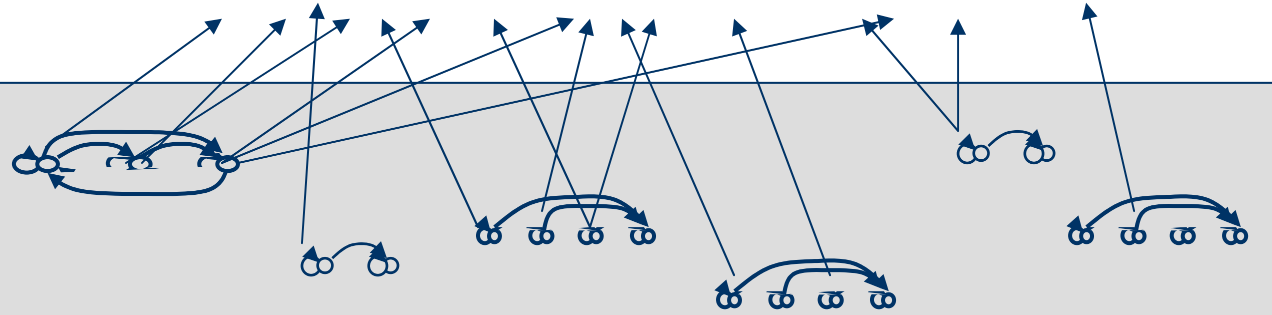
HDESM Example (HMM):



3 states + transition probabilities
n observable events: a,b,c,d,e,…
Pr(state | observable event) given/known

Observed event sequence:

….abcbbbaaaababbabcccbdddbebdbabcbabe….



Catalog of Processes

Which combination of which process models “best” accounts for the observations?
 Events not associated with a known process are “ANOMALIES”.

An analogy....

What does

hbeolnjouolor

mean?

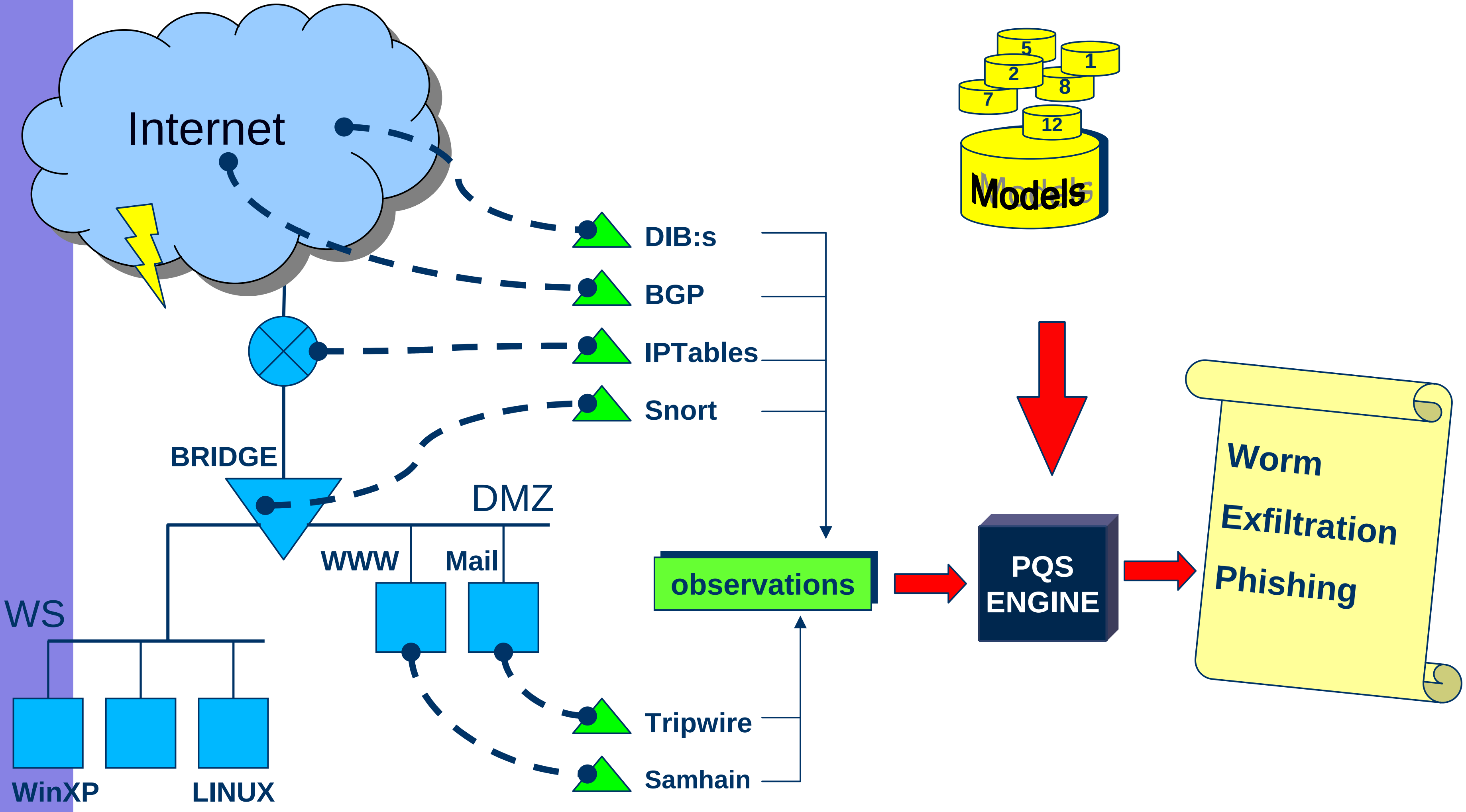
Events are: hbeolnjouolor

Models = French + English words (+ grammars!)

hbeolnjoulor = hello + bonjour

Intermediate hypotheses include tracks: ho + be

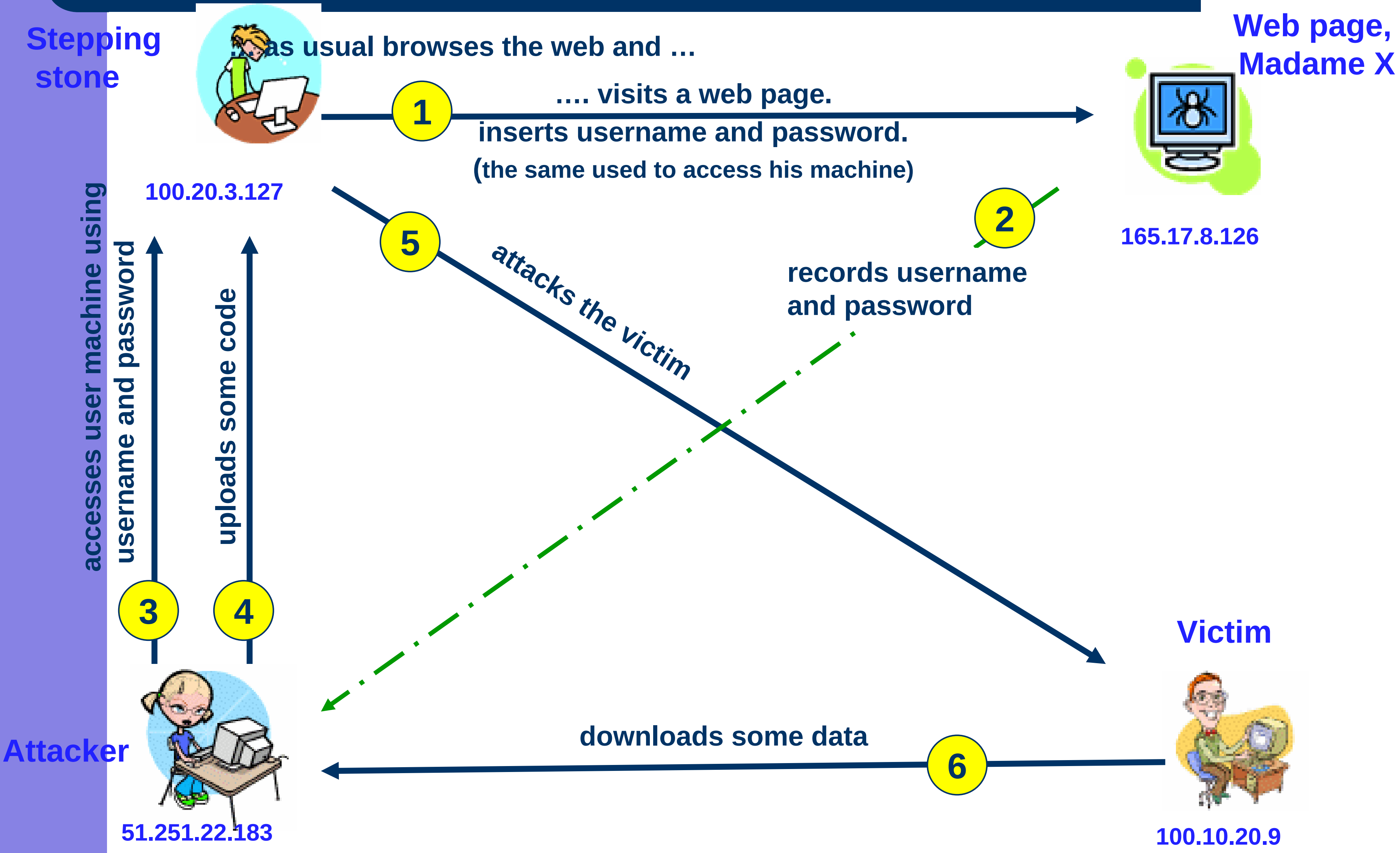
PQS in Computer Security



Outline

1. Motivation and Terminology
2. Process Query System (PQS) Approach
3. Detection of a complex attack
4. Conclusion and Acknowledgments

Complex Phishing Attack Steps



Stepping stone

… as usual browses the web and …

1 …. visits a web page.

inserts username and password.

(the same used to access his machine)

Web page, Madame X

165.17.8.126

100.20.3.127

2 records username and password

5 attacks the victim

3 accesses user machine using username and password

4 uploads some code

Attacker

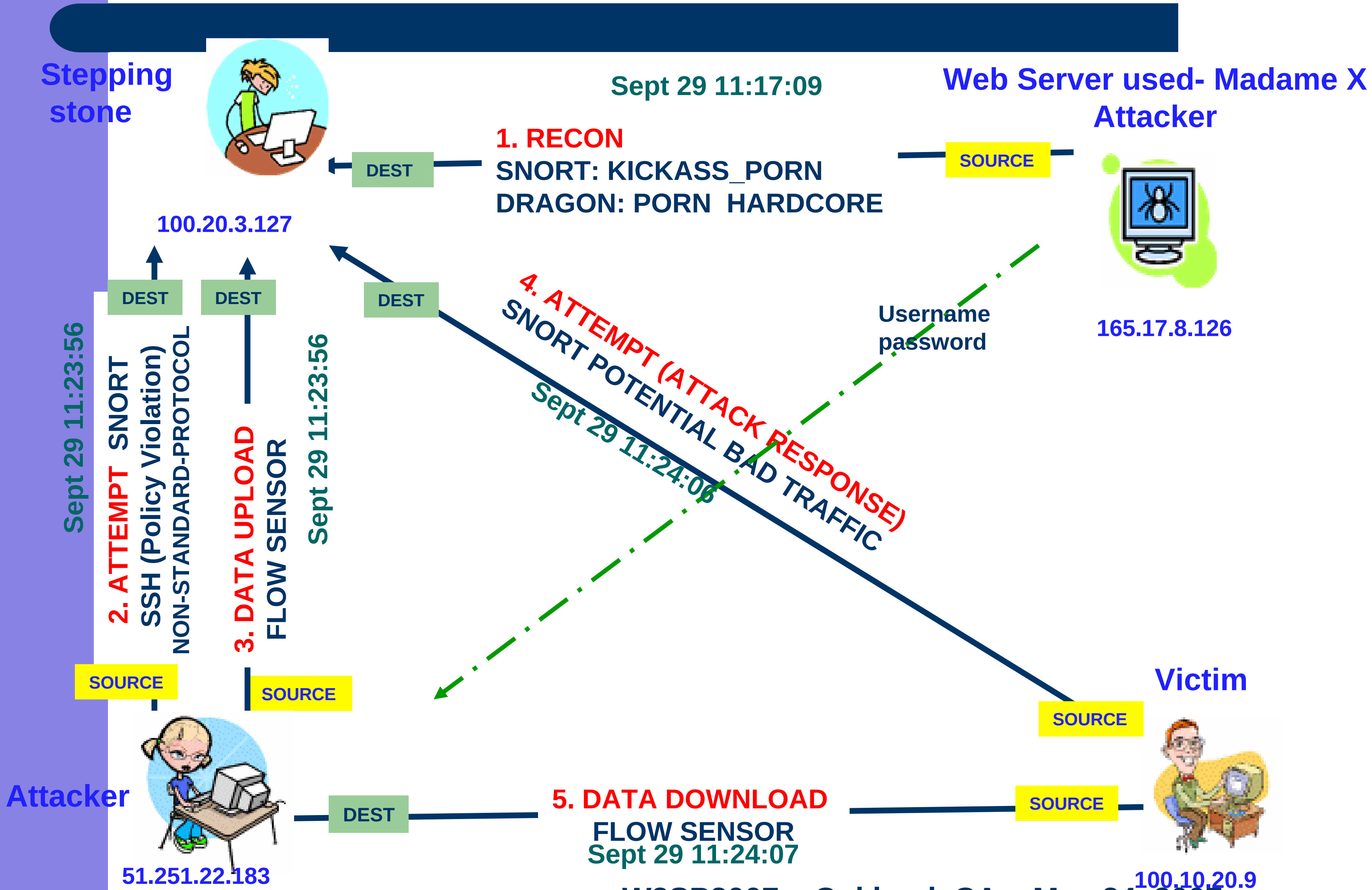
51.251.22.183

Victim

100.10.20.9

6 downloads some data

Complex Phishing Attack Observables



Stepping stone

100.20.3.127

Web Server used- Madame X Attacker

165.17.8.126

Attacker

51.251.22.183

Victim

100.10.20.9

Sept 29 11:17:09

1. RECON
SNORT: KICKASS_PORN
DRAGON: PORN HARDCORE

Sept 29 11:23:56
2. ATTEMPT SNORT
SSH (Policy Violation)
NON-STANDARD-PROTOCOL

3. DATA UPLOAD
FLOW SENSOR
Sept 29 11:23:56

4. ATTEMPT (ATTACK RESPONSE)
SNORT POTENTIAL BAD TRAFFIC
Sept 29 11:24:06

Username
password

5. DATA DOWNLOAD
FLOW SENSOR
Sept 29 11:24:07

Flow Sensor

- Based on the *libpcap* interface for packet capturing.
- Packets with the same source IP, destination IP, source port, destination port, protocol are aggregated into the same flow.

  - Timestamp of the last packet
  - # packets from Source to Destination
  - # packets from Destination to Source
  - # bytes from Source to Destination
  - # bytes from Destination to Source
  - Array containing delays in microseconds between packets in the flow

We did not use *Netflow* only because it does not have all the fields that we need.

Two Models Based on the Flow Sensor

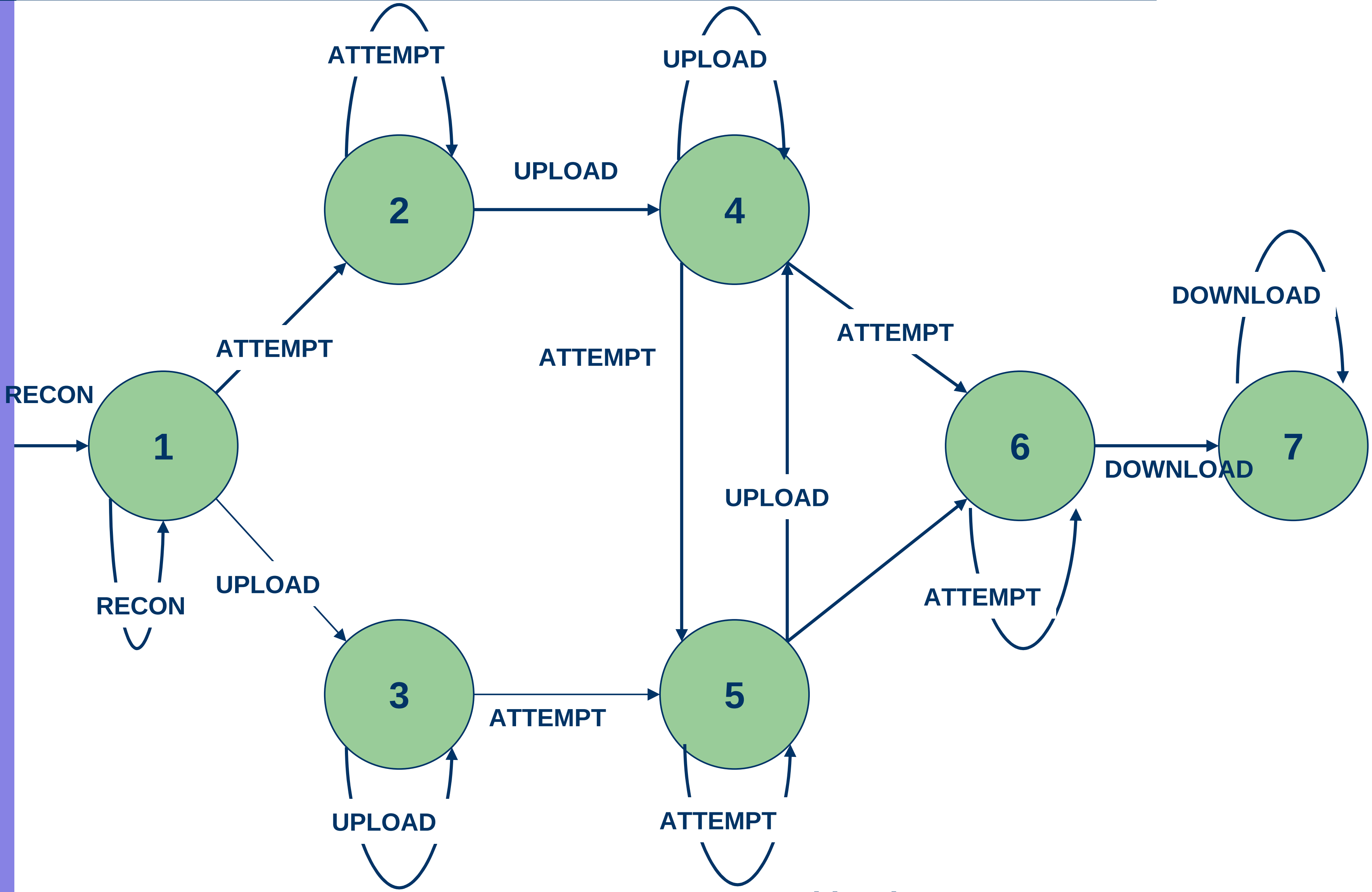
Low and Slow UPLOAD

Volume	Packets	Duration	Balance	Percentage
Tiny: 1-128b Small: 128b-1Kb	4:10-99 5: 100-999 6: > 1000	4: 1000-10000 s 5: 10000-100000 s 6: > 100000 s	Out	>80

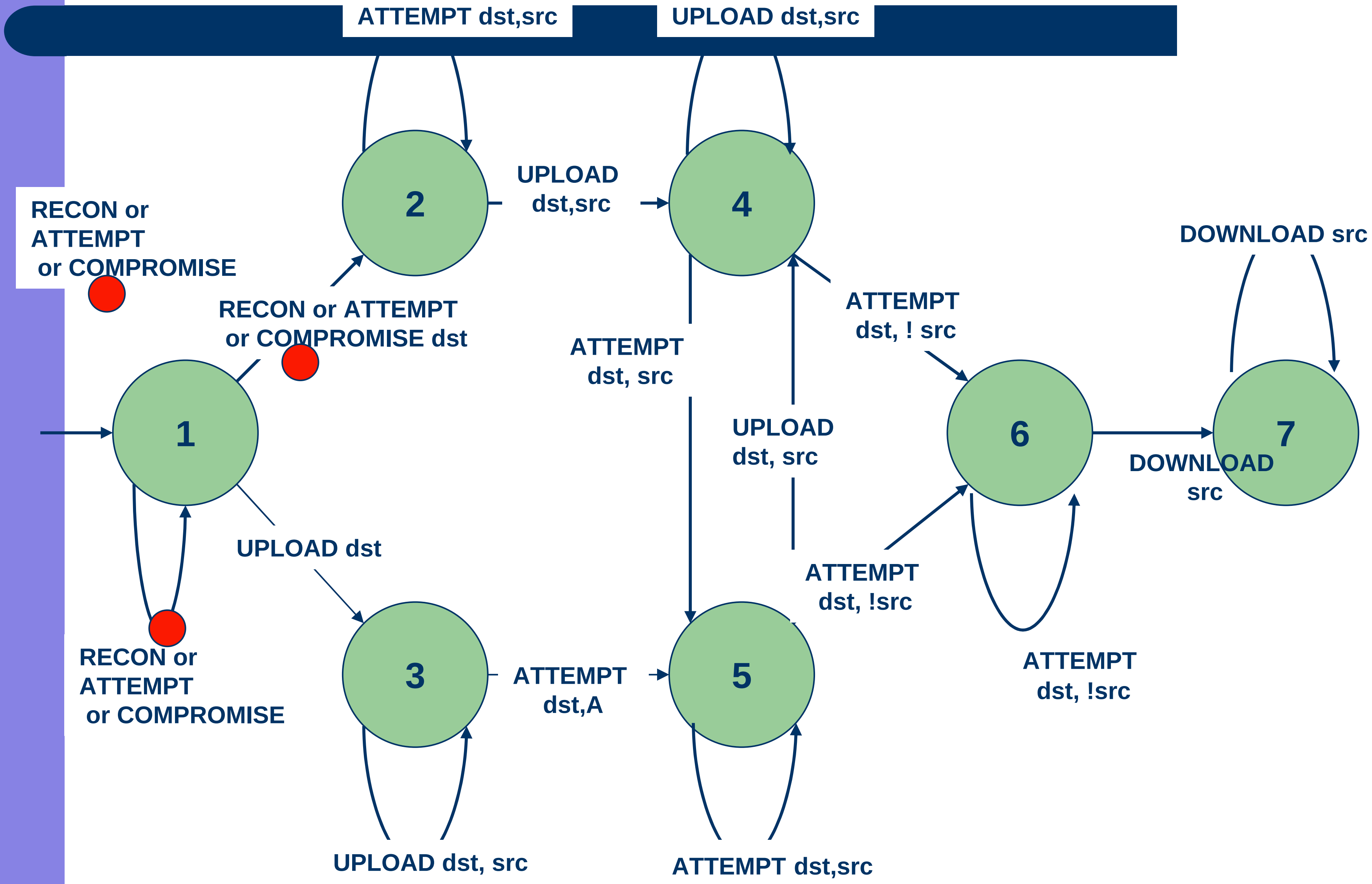
UPLOAD

Volume	Packets	Duration	Balance	Percentage
Tiny: 1-128b Small: 128b-1Kb Medium: 1Kb-100Kb Large: > 100Kb	1: one packet 2: two pckts 3: 3-9 4: 10-99 5: 100-999 6: > 1000	0: < 1 s 1: 1-10 s 2: 10-100 s 3: 100-1000 s 4: 1000-10000 s 5: 10000-100000 s 6: > 100000 s	Out	>80

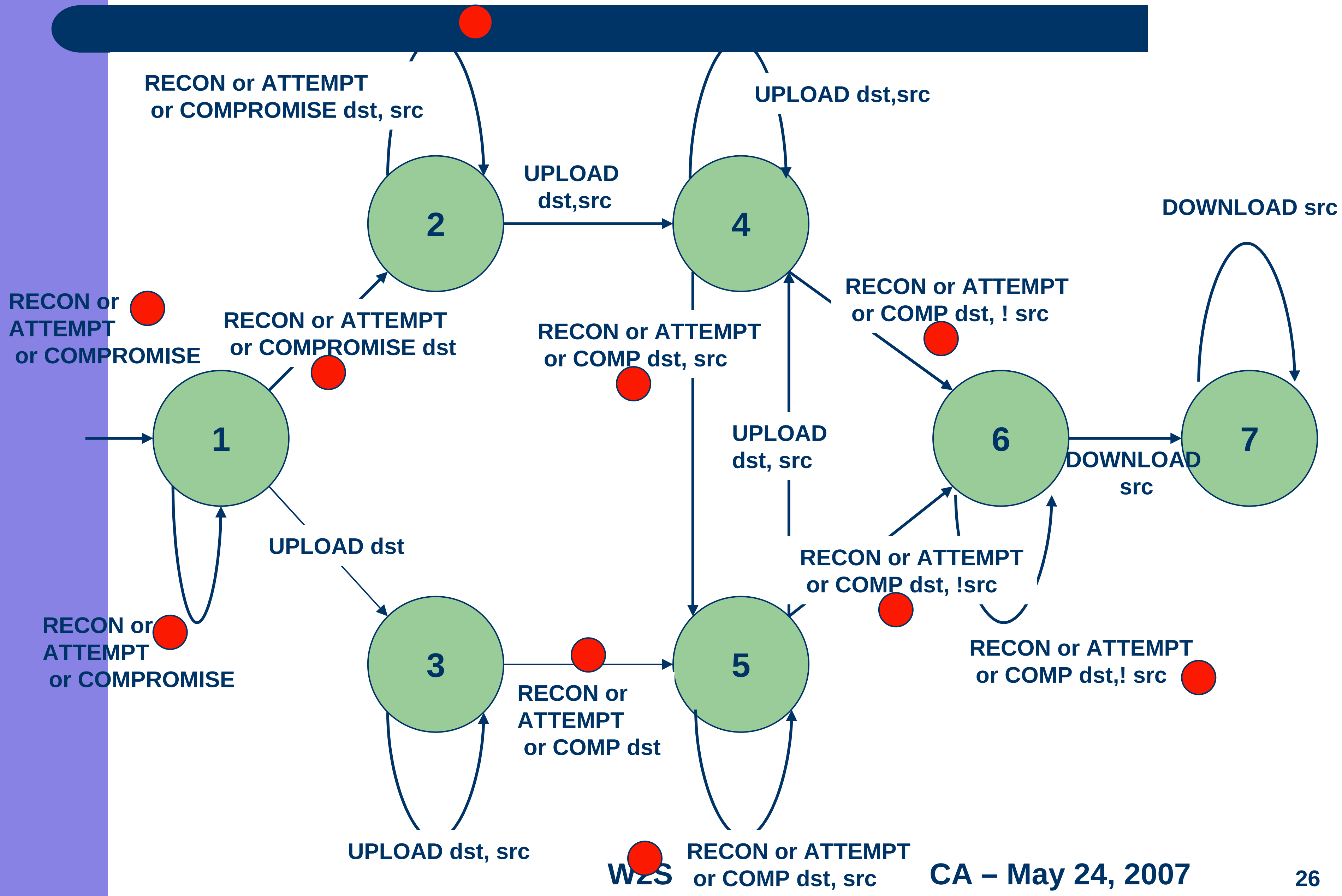
Phishing Attack Model 1 – very specific



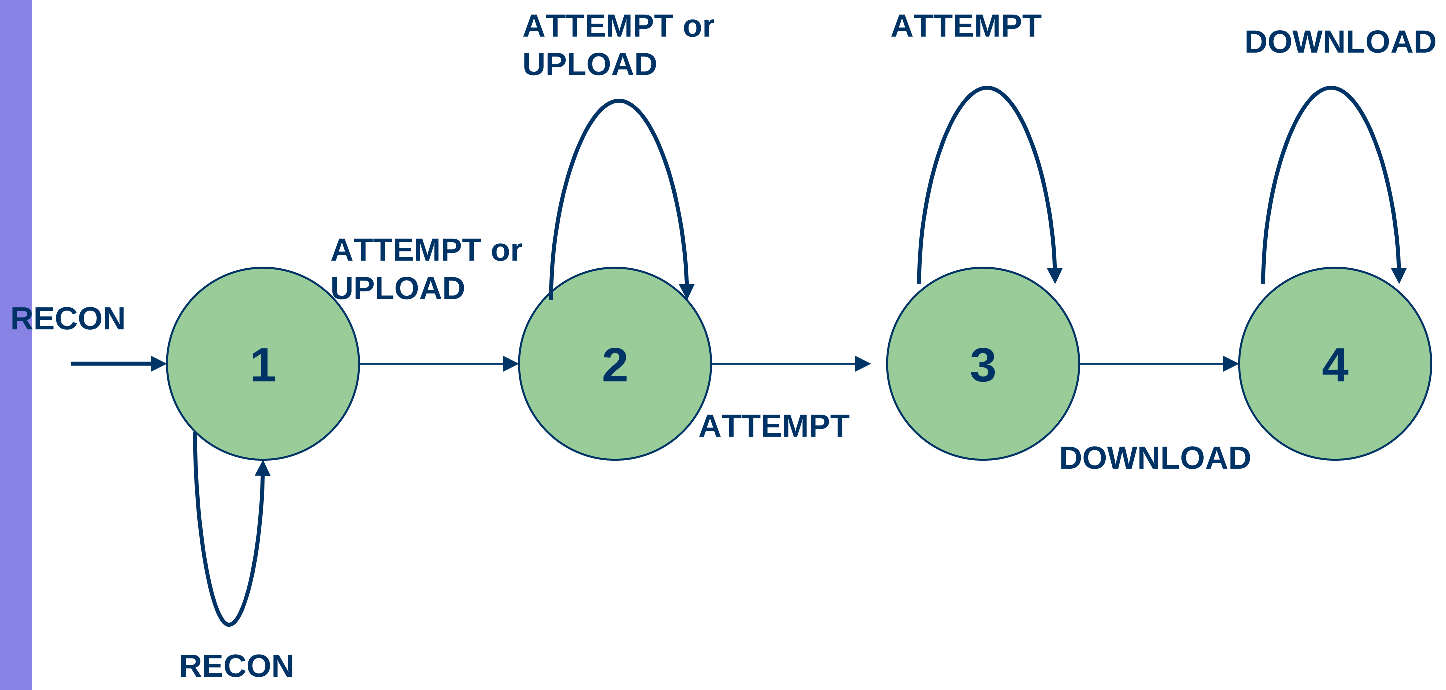
Phishing Attack Model 2 – less specific



Phishing Attack Model 3 – more general



Phishing Attack Model 3 – Most general



Stricter models reduce false positives, **but** less strict models can detect unknown attack sequences

Outline

1. Motivation and Terminology
2. Process Query System (PQS) Approach
3. Detection of a complex attack
4. Conclusion and Acknowledgments

Contribution

- Identification of a new generation of threats
- Need for new paradigms of combining alerts (observations)
- Process Query System (PQS) based approaches to detect complex attacks and covert channels
- Need of reducing the gap between user perception and what technology means (maybe explicit information about the real status of the system).

Many thanks to professor George Cybenko (Thayer School of Engineering at Dartmouth College) and professor Shankar Sastry (EECS, UC Berkeley).

agiani@eecs.berkeley.edu