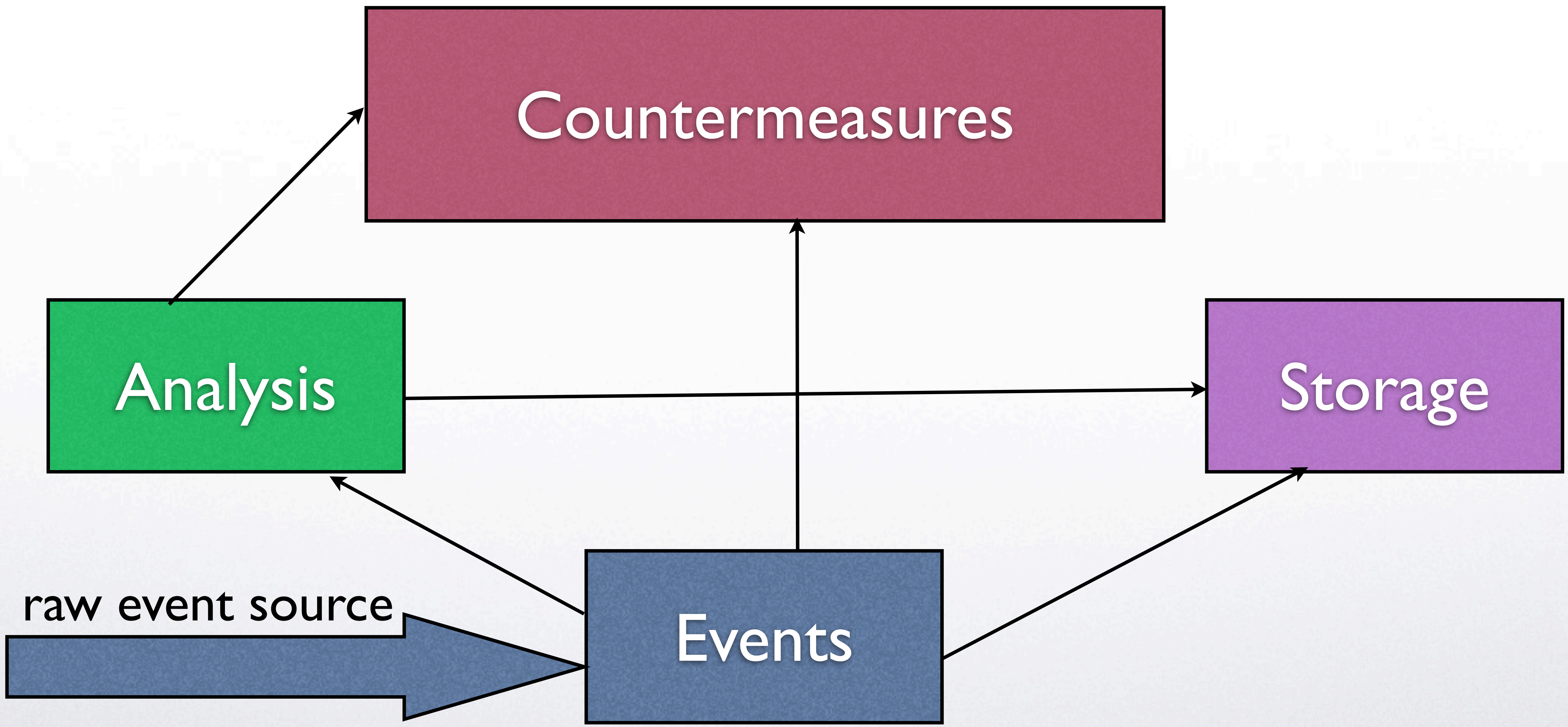
Intrusion Detection

Aggelos Kiayias

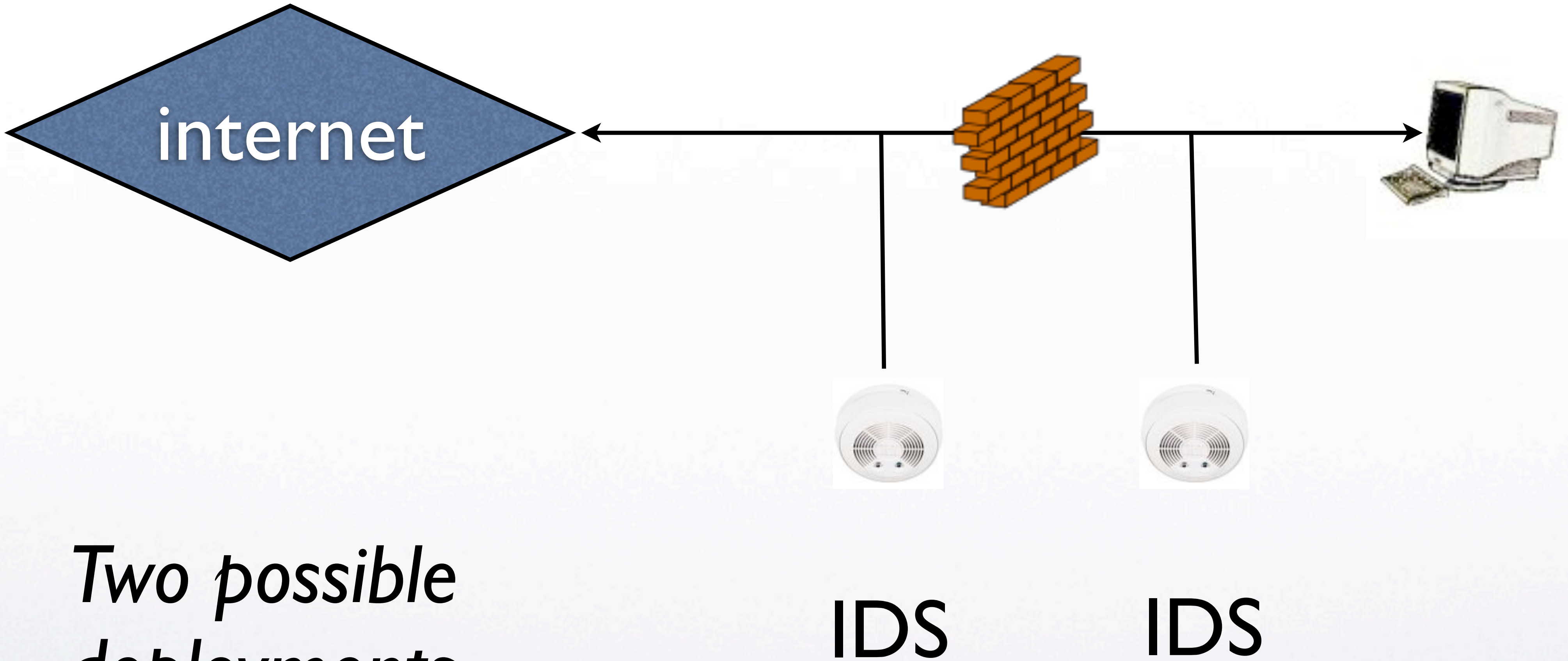
IDS

- your system's smoke detector.

Intrusion Detection System



IDS & Firewall



Two possible deployments based on a wiretap

Functions

- monitoring users and system activity
- auditing system configuration for vulnerabilities and misconfigurations
- assessing the integrity of critical system and data files.
- recognizing known attack patterns in system activity.
- identifying abnormal activity by statistical analysis.
- create & manage audit trail.
- installing and operating traps to record information about intruders.

Types of IDS

with respect to detection approach

- Signature based.
  - pattern matching *triggers action*.
- Anomaly based.
  - building on a model of acceptable behavior. Deviation from acceptable behavior *triggers action*.

Types of IDS

with respect to deployment position

- Network-based.
  - running on the network as a packet sniffer.
- Host-based.
  - running on the host as an event monitor.

Network Based IDS

- Operates on raw network data
  - trap packets from a network interface operating in promiscuous mode.
  - scan packets for known attack patterns (e.g., DoS packets, shellcodes, connection attempts to inappropriate ports, etc.).
- Once an attack pattern is discovered a countermeasure can be activated.

Host Based IDS

- Based on system logs and system events.
  - E.g., trigger action when certain files get opened or modified, or in creation/deletion of user accounts, root shells, etc.
  - reaction can be near real time for the system under attack.

Host vs. Network

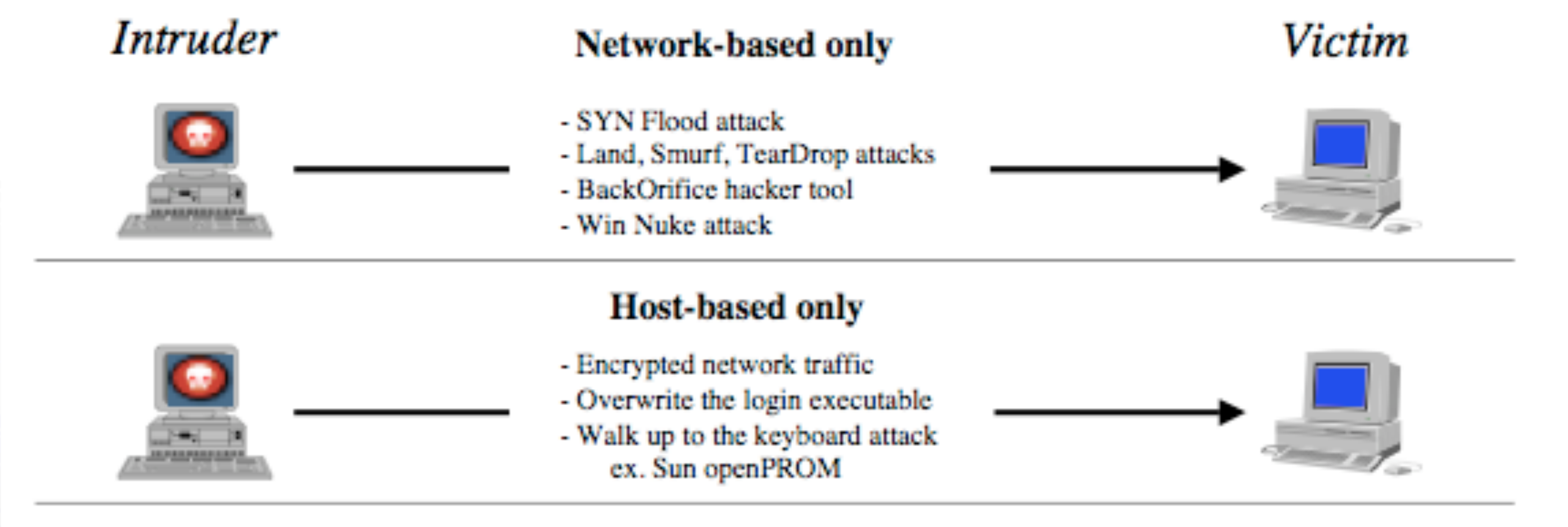


figure from:Intrusion Detection Systems, B.Laing, http://www.snort.org/docs/iss-placement.pdf

Land Attack: SYN with sourceIP=targetIP
Smurf Attack: PING with spoofed sourceIP=victimIP sent to broadcast address
Teardrop Attack: bad fragmentation invalid offset.
Winnuke attack: using TCP feature that was not handled properly by MS TCP/IP.

Host vs. Network, II

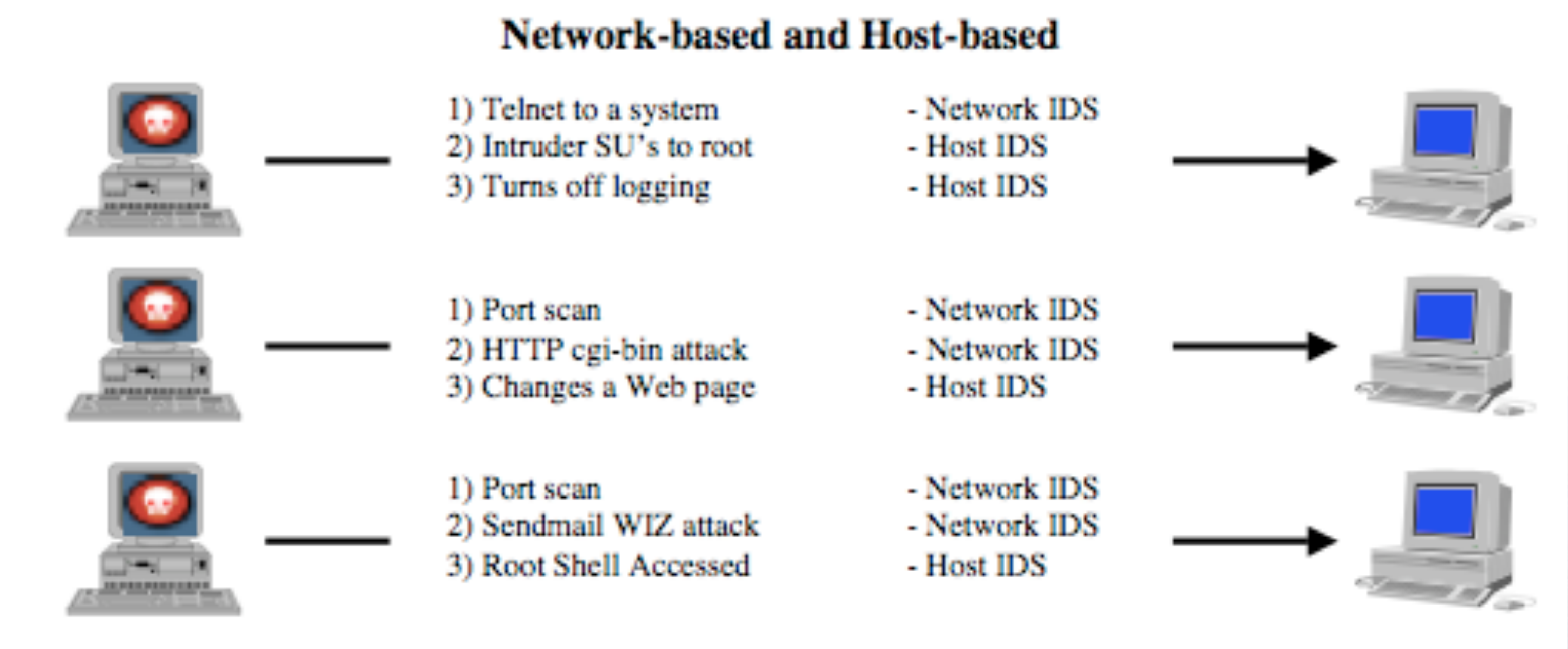


figure from:Intrusion Detection Systems, B.Laing, http://www.snort.org/docs/iss-placement.pdf

Host vs. Network, III

- Network-based: low cost, employs packet analysis, can work for any system, detects unsuccessful attacks, leaves behind proofs of attack attempt and malicious intent.
- Host-based: less false positives, non-network attacks (walk-in), network deployment independence.

Signature based IDS

- Better detection capability (lower false positive rate).
- Relies on rules that are built based on known attacks.
- Possibility of evasion by slightly modifying attack traffic.

Anomaly Based IDS

- Lower detection capability but potentially captures even unknown attacks.
- Based on a statistical model to figure out what is anomalous to begin with.
- May yield high false positive rate.

IDS Systems

- **Tripwire**: host-based IDS. File integrity checking and reporting. http://sourceforge.net/projects/tripwire/
- **Snort**: network-based IDS. Real-time traffic analysis. http://www.snort.org/
- **Bro**: network analysis framework. Real-time traffic analysis. http://www.bro.org/

Honeypots

- Hosts that are left vulnerable on purpose.
- May be set to contain seemingly valuable information.
- Can be used to log attacker (IP address, keystrokes, software attempted to install).
- Feed results into your IDS/Firewall system.

Darknet Server

- Receives all traffic that is directed to unused IP space of a subnetwork.
- Can be combined with a honeypot.

The False Positive Problem

figure from: Aesopus; Steinhöwel, Heinrich; Brant, Sebastian: Esopi appologi sive mythologi: cum quibusdam carminum et fabularum additionibus Sebastiani Brant.

Conditional probability

A event

Ω space

Probability $\mathbf{Prob}[A] = \dfrac{\#A}{\#\Omega}, A \subseteq \Omega$

Conditional Probability

$$\mathbf{Prob}[A \mid B] = \frac{\mathbf{Prob}[A \cap B]}{\mathbf{Prob}[B]} = \frac{\#(A \cap B)}{\#B}$$

Total Probability Theorem E_i partition Ω

$$\mathbf{Prob}[A] = \sum_i \mathbf{Prob}[A \cap E_i] = \sum_i \mathbf{Prob}[A \mid E_i]\mathbf{Prob}[E_i]$$

Bayes Theorem

$$\mathbf{Prob}[E_1 \mid A] = \frac{\mathbf{Prob}[A \mid E_1]\mathbf{Prob}[E_1]}{\mathbf{Prob}[A]}$$

where the denominator can be expressed as

$$\mathbf{Prob}[A] = \sum_i \mathbf{Prob}[A \mid E_i]\mathbf{Prob}[E_i]$$

Hypothesis testing: what is the probability that the hypothesis E_1 is true given evidence A

Medical Test Paradox

- Suppose: A laboratory test for a disease is 98% accurate and has a 3% probability of a false positive.
- The prevalence of the disease is 1% in the population.
- Suppose you test positive. What are the chances you are sick?

Relation to IDS

- A test for a type of malicious traffic is 98% accurate and has a 3% probability of a false positive.
- From all traffic only 1% is malicious.
- What is the probability that you are under attack given that the IDS triggers the alarm?

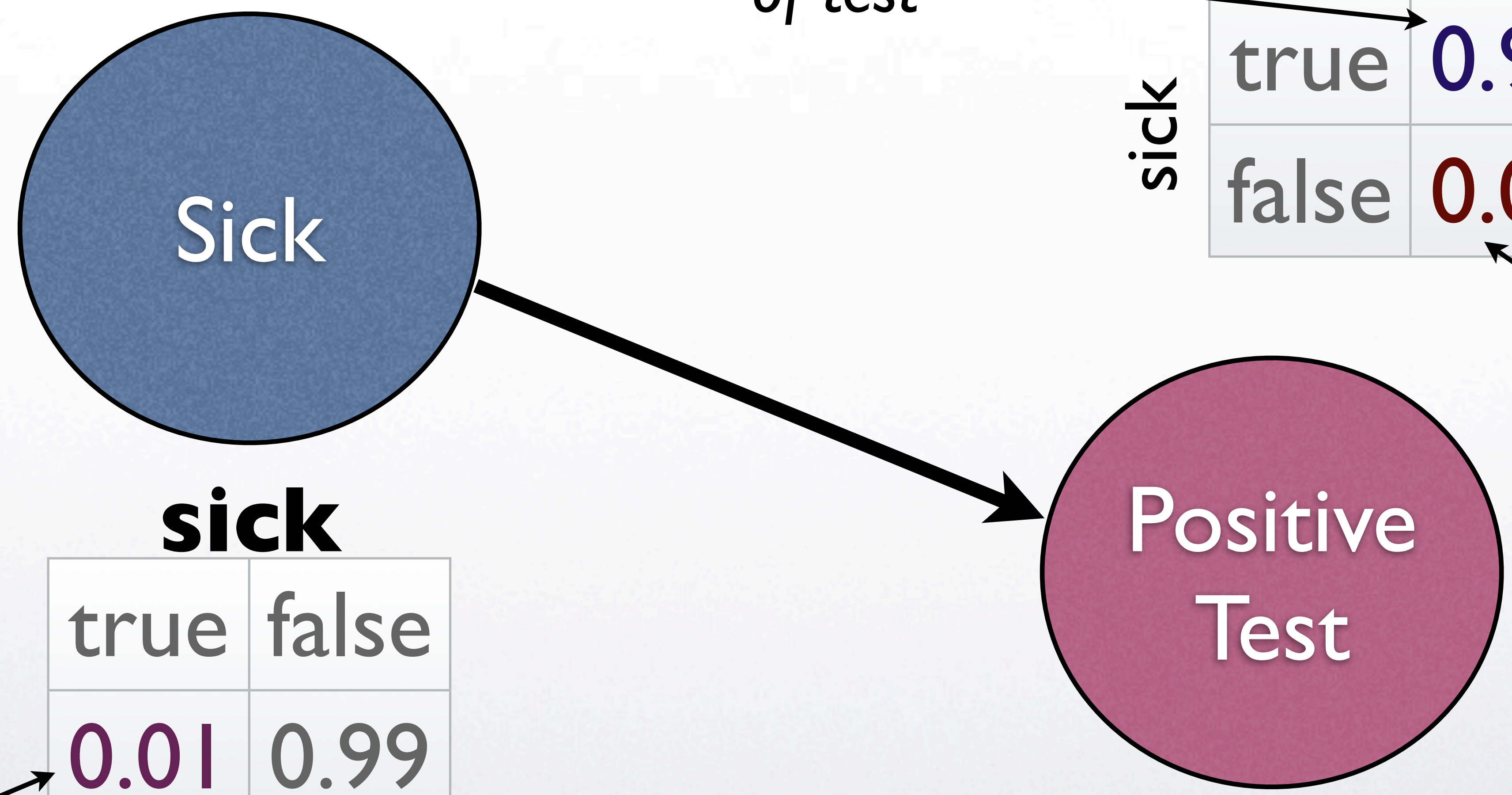
Causality

false negative

accuracy of test

		test positive	
		true	false
sick	true	0.98	0.02
	false	0.03	0.97

false positive



sick

true	false
0.01	0.99

prevalence

Calculate Prob[Sick|Positive Test]=?

Bayesian Testing

seq sequence of packets = random variable over your traffic

AttackType a particular type of attack

AttackPattern a certain pattern that we can test the traffic on and may potentially suggest an attack.

hypothesis

the test employed by the IDS is positive

$$\mathbf{Prob}[\text{seq} \in \text{AttackType} \mid \text{seq} \in \text{AttackPattern}] =$$

probability estimated through training (accuracy)

estimated parameter

$$= \frac{\mathbf{Prob}[\text{seq} \in \text{AttackPattern} \mid \text{seq} \in \text{AttackType}]\mathbf{Prob}[\text{seq} \in \text{AttackType}]}{\mathbf{Prob}[\text{seq} \in \text{AttackPattern}]}$$

Bayesian Testing, II

for the denominator

$$\mathbf{Prob}[\text{seq} \in \text{AttackPattern}] =$$

accuracy

parameter

$$\mathbf{Prob}[\text{seq} \in \text{AttackPattern} \mid \text{seq} \in \text{AttackType}]\mathbf{Prob}[\text{seq} \in \text{AttackType}]+$$

false positive (also through training)

parameter

$$+\mathbf{Prob}[\text{seq} \in \text{AttackPattern} \mid \text{seq} \notin \text{AttackType}]\mathbf{Prob}[\text{seq} \notin \text{AttackType}]$$

Based on the above we obtain the probability of the hypothesis being true given the observation

Bayesian Testing, III

- Decision based on a single test insufficient.
- Employing more tests:

$$\mathbf{Prob}[\mathbf{S} \mid \mathbf{T}_1 \wedge \ldots \wedge \mathbf{T}_n] = ?$$

Using Bayes Theorem:

$$= \frac{\mathbf{Prob}[\mathbf{T}_1 \wedge \ldots \wedge \mathbf{T}_n \mid \mathbf{S}] \cdot \mathbf{Prob}[\mathbf{S}]}{\mathbf{Prob}[\mathbf{T}_1 \wedge \ldots \wedge \mathbf{T}_n]}$$

Bayesian Testing IV

$$\mathbf{Prob}[\mathbf{T}_1 \wedge \ldots \wedge \mathbf{T}_n \mid \mathbf{S}] = \mathbf{Prob}[\mathbf{T}_1 \mid \mathbf{S}] \cdot$$
$$\mathbf{Prob}[\mathbf{T}_2 \mid \mathbf{S} \wedge \mathbf{T}_1] \cdot$$
$$\mathbf{Prob}[\mathbf{T}_3 \mid \mathbf{S} \wedge \mathbf{T}_1 \wedge \mathbf{T}_2] \cdot$$
$$\mathbf{Prob}[\mathbf{T}_4 \mid \mathbf{S} \wedge \mathbf{T}_1 \wedge \mathbf{T}_2 \wedge \mathbf{T}_3] \cdot$$
$$\ldots$$

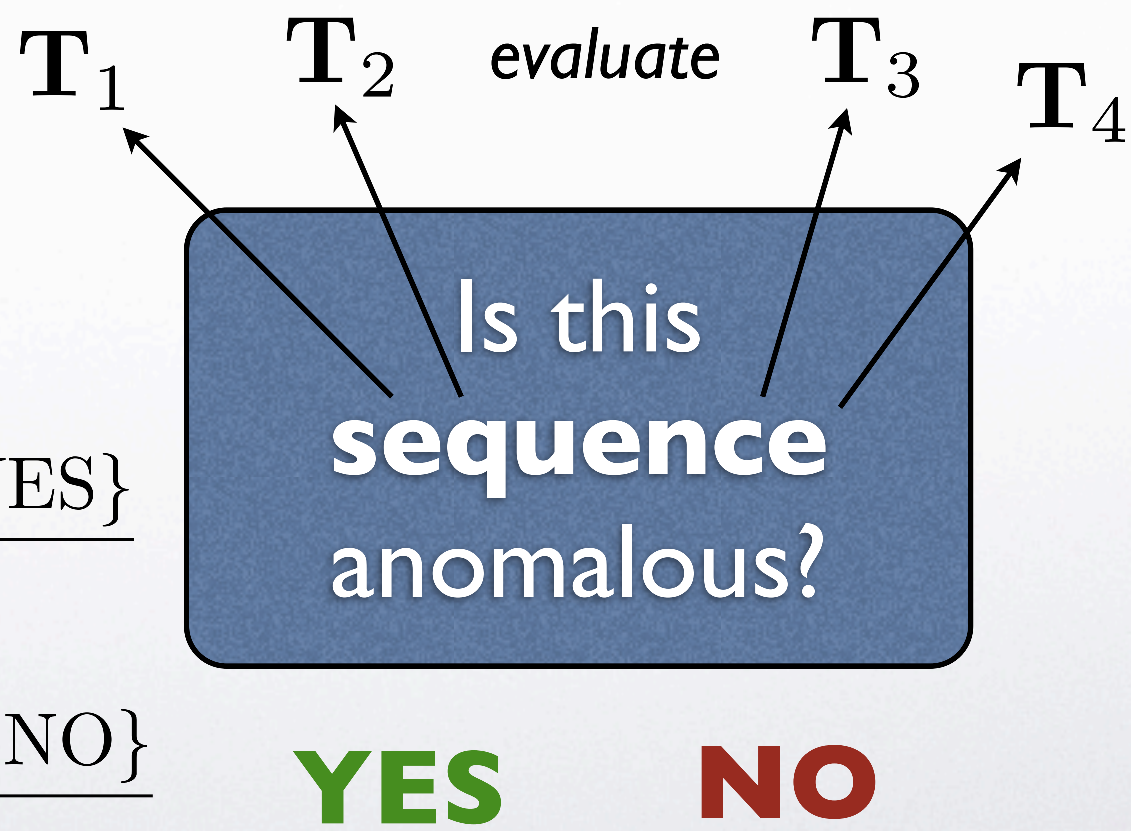
The naive approach: assume all tests are independent variables.

$$= \prod_{i=1}^{n} \mathbf{Prob}[\mathbf{T}_i \mid \mathbf{S}]$$

Training

- Training phase: estimation of $\mathbf{Prob}[\mathbf{T}_i \mid \mathbf{S}]$ $\mathbf{Prob}[\mathbf{T}_i \mid \neg\mathbf{S}]$

Example:
Human assisted.



accuracy

$$\mathbf{Prob}[\mathbf{T}_i \mid \mathbf{S}] = \frac{\#\{\mathbf{T}_i = \text{true} \mid \text{YES}\}}{\text{total YES}}$$

false-positive

$$\mathbf{Prob}[\mathbf{T}_i \mid \neg\mathbf{S}] = \frac{\#\{\mathbf{T}_i = \text{true} \mid \text{NO}\}}{\text{total NO}}$$

Bayesian Inference

- A well trained Bayesian inference system can be very successful in properly classifying input sequences as anomalous or not.
- Gives rise to anomaly based IDS with good overall false positive rates.