



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architetture

Posizionamento
sensori

Sicurezza dei sistemi e delle reti¹

Mattia Monga

Dip. di Informatica
Università degli Studi di Milano, Italia
mattia.monga@unimi.it

a.a. 2015/16

¹ © 2011–15 M. Monga. Creative Commons Attribuzione — Condividi allo stesso modo 4.0 Internazionale. <http://creativecommons.org/licenses/by-sa/4.0/deed.it>. Derivato con permesso da © 2010 M. Cremonini.



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architetture

Posizionamento
sensori

Lezione IX: Falsi allarmi



In generale in tutti gli IDS (misuse e anomaly) occorre bilanciare

falsi negativi attacchi non rilevati

falsi positivi attacchi rilevati corrispondenti a situazioni normali



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

- Quanto piú la rilevazione è **specific**a (es. firme molto dettagliate) tanto piú aumenta il carico computazionale e la rilevazione diventa sensibile a variazioni dell'evento analizzato.
- Quanto piú la rilevazione si fa **lasca** (es. firme generiche) tanto piú il carico computazionale cala, la rilevazione dell'evento analizzato risulta poco influenzata da varianti ma tanto piú vengono rilevati eventi simili ma non pericolosi.

Relazione fra FP e FN



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architettonici

Posizionamento
sensori

FP e FN risultano correlati inversamente: agendo per diminuire l'una, tipicamente l'altra aumenta.

Il problema si ripropone in moltissime discipline (information retrieval, farmacologia, ...): ogni volta che si ha una decisione binaria (test)

	positivo	negativo
attacco	TP	FN
non attacco	FP	TN

$$TP + TN + FP + FN = totale$$



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architetture

Posizionamento
sensori

FP Type I error, falso allarme

FN Type II error, miss

sensibilità del test, recall, hit rate, TPR $\frac{TP}{TP+FN}$

specificità del test $\frac{TN}{TN+FP}$

accuratezza del test $\frac{TP+TN}{totale}$

precisione del test $\frac{TP}{TP+FP}$

FPR $\frac{FP}{TN+FP} = 1 - \text{specificità}$



Quali regole di IDS hanno lavorato meglio?

Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architettureali

Posizionamento
sensori

A	allarme	$\neg$ allarme	B	allarme	$\neg$ allarme
attacco	TP=63	FN=37	attacco	TP=77	FN=23
$\neg$ attacco	FP=28	TN=72	$\neg$ attacco	FP=77	TN=23
C	allarme	$\neg$ allarme	D	allarme	$\neg$ allarme
attacco	TP=24	FN=76	attacco	TP=76	FN=24
$\neg$ attacco	FP=88	TN=12	$\neg$ attacco	FP=12	TN=88



Quali regole di IDS hanno lavorato meglio?

A	allarme	$\neg$ allarme	B più sensibile	allarme	$\neg$ allarme
attacco	TP=63	FN=37	attacco	TP=77	FN=23
$\neg$ attacco	FP=28	TN=72	$\neg$ attacco	FP=77	TN=23
C	allarme	$\neg$ allarme			
attacco	TP=24	FN=76			
$\neg$ attacco	FP=88	TN=12			
D più specifico, accurato, preciso	allarme	$\neg$ allarme			
attacco	TP=76	FN=24			
$\neg$ attacco	FP=12	TN=88			

Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi
Teorema di
Bayes

Aspetti
architeturali
Posizionamento
sensori



Quali regole di IDS hanno lavorato meglio?

A (FPR:0.28,TPR:0.63)	allarme	¬allarme
attacco	TP=63	FN=37
¬attacco	FP=28	TN=72

B (0.77,0.77)	allarme	¬allarme
attacco	TP=77	FN=23
¬attacco	FP=77	TN=23

C (0.88,0.24)	allarme	¬allarme
attacco	TP=24	FN=76
¬attacco	FP=88	TN=12

D (0.12,0.76)	allarme	¬allarme
attacco	TP=76	FN=24
¬attacco	FP=12	TN=88

Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

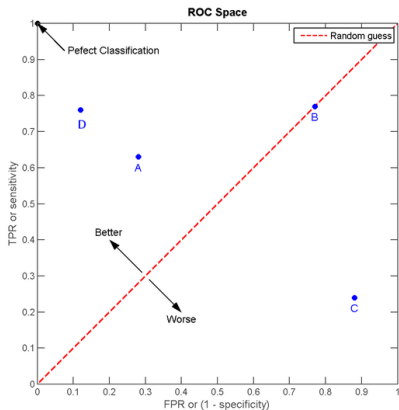
I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

receiver operating characteristic (ROC): sensibilità vs. tasso dei falsi positivi



Sicurezza delle reti

Monga

Falsi allarmi

Rilevamento delle intrusioni

I falsi allarmi

Teorema di Bayes

Aspetti architetturelli

Posizionamento sensori

E un insieme di regole?



Sicurezza delle
reti

Monga

Falsi allarmi

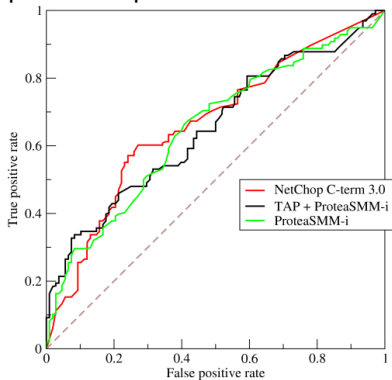
Rilevamento
delle intrusioni

I falsi allarmi
Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

E come valutare l'efficacia di un insieme di regole per tutti i parametri possibili? A volte si usa l'Area under curve.





Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

Il problema degli IDS sono i falsi allarmi e le mancate segnalazioni

- la qualità di un IDS sta nella relazione fra FP e FN
- per valutarla ci si serve di ROC e AUC

I falsi allarmi



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architettonici

Posizionamento
sensori

Gli strumenti di analisi come ROC permettono di valutare l'efficacia *ex-post*, valutando il peso di FP e FN in un determinato contesto sperimentale.

Un IDS genera migliaia di allarmi al giorno: qual è la probabilità che un allarme sia davvero relativo a un attacco?

L'intuito inganna perché dipende in maniera complessa dalla
probabilità a priori di un attacco.



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architetture

Posizionamento
sensori

La probabilità che una donna sviluppi un cancro al seno è 0.8%. Se una donna **ha** il cancro al seno, la probabilità che il suo mammogramma sia positivo è 90%; se **non ha** il cancro al seno, c'è comunque una probabilità del 7% che il mammogramma sia positivo. Se il mammogramma di una donna è positivo, qual è la probabilità che abbia effettivamente il cancro al seno?



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

La probabilità che una donna sviluppi un cancro al seno è 0.8%. Se una donna **ha** il cancro al seno, la probabilità che il suo mammogramma sia positivo è 90%; se **non ha** il cancro al seno, c'è comunque una probabilità del 7% che il mammogramma sia positivo. **Se il mammogramma di una donna è positivo, qual è la probabilità che abbia effettivamente il cancro al seno?**

Esempio da "Quando i numeri ingannano", di G. Gigerenzer: studi su medici mostrano che la risposta più frequente al problema così posto è 90%.

Esempio: soluzione



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

- Su 1000 donne, 992 sono sane e 8 malate.
- Delle 8 malate, il 90% ($\simeq 7$) risulteranno positive e il 10% negative ($\simeq 1$).
- Delle 992 sane, il 7% ($\simeq 70$) risulteranno positive e il 93% negative ($\simeq 922$).
- Le positive saranno quindi $7 + 70 = 77$, delle quali sono malate 7: la probabilità che un mammogramma positivo sia indice di malattia è quindi $\frac{7}{77} = 9\%$

Teorema di Bayes



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

**Teorema di
Bayes**

Aspetti
architetture

Posizionamento
sensori

Teorema di Bayes

$$\Pr(\text{attacco}|\text{allarme}) = \frac{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco})}{\Pr(\text{allarme})}$$

Teorema di Bayes



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

Teorema di Bayes

$$\Pr(\text{attacco}|\text{allarme}) = \frac{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco})}{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco}) + \Pr(\text{allarme}|\neg\text{attacco}) \cdot \Pr(\neg\text{attacco})}$$

Teorema di Bayes



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori

Teorema di Bayes

$$\Pr(\text{attacco}|\text{allarme}) = \frac{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco})}{\Pr(\text{allarme})}$$
$$\left\{ \frac{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco})}{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco}) + \Pr(\text{allarme}|\neg\text{attacco}) \cdot \Pr(\neg\text{attacco})} \right.$$
$$\left\{ \frac{\frac{TP}{TP+FN} \cdot \Pr(\text{attacco})}{\frac{TP}{TP+FN} \cdot \Pr(\text{attacco}) + \frac{FP}{FP+TN} \cdot \Pr(\neg\text{attacco})} \right.$$



Teorema di Bayes

Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architetture

Posizionamento
sensori

Teorema di Bayes

$$\Pr(\text{attacco}|\text{allarme}) = \frac{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco})}{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco}) + \Pr(\text{allarme}|\neg\text{attacco}) \cdot \Pr(\neg\text{attacco})}$$
$$\left\{ \begin{array}{l} \frac{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco})}{\Pr(\text{allarme}|\text{attacco}) \cdot \Pr(\text{attacco}) + \Pr(\text{allarme}|\neg\text{attacco}) \cdot \Pr(\neg\text{attacco})} \\ \frac{\frac{TP}{TP+FN} \cdot \Pr(\text{attacco})}{\frac{TP}{TP+FN} \cdot \Pr(\text{attacco}) + \frac{FP}{FP+TN} \cdot \Pr(\neg\text{attacco})} \end{array} \right.$$

Per calcolare la probabilità di un allarme veritiero occorre sempre stimare **la probabilità a priori di un attacco**, che è spesso (fortunatamente!) piuttosto bassa: **i falsi allarmi** sono inevitabilmente comuni, a meno di avere un IDS straordinariamente preciso o asset particolarmente appetibili.



Riprendiamo il migliore IDS esaminato con la curva ROC

D (0.12,0.76)	allarme	¬allarme
attacco	TP=76	FN=24
¬attacco	FP=12	TN=88

- Nell'esperimento:

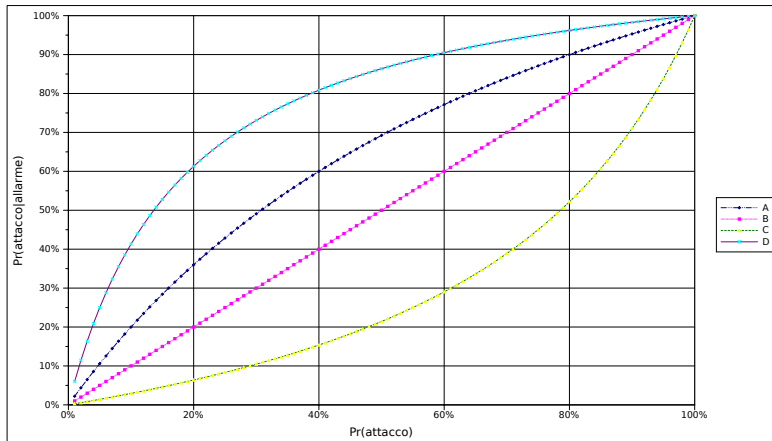
$$\Pr(\text{attacco}) = \frac{76+24}{76+24+12+88} = 50\%$$

$$\rightsquigarrow \Pr(\text{attacco}|\text{allarme}) = 86\%$$

- Attacchi poco frequenti

$$\Pr(\text{attacco}) = 1\%$$

$$\rightsquigarrow \Pr(\text{attacco}|\text{allarme}) = 6\%$$





Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architettonici

Posizionamento
sensori

- L'amministratore di un IDS è interessato a stimare la probabilità che un allarme sia davvero relativo a un attacco.
- dipende però dalla probabilità *a priori* di un attacco.



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architetturali

Posizionamento
sensori

Un NIDS **complementa** altre soluzioni, con un'architettura a diversi livelli (**defense in-depth**).

Importanti aspetti architetturali:

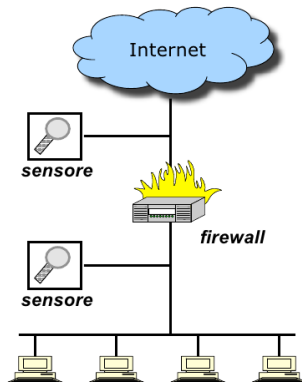
- Quanti sensori installare nella rete
 - costi e complessità di gestione
- Dove installarli
 - quantità vs. ridondanza di informazioni
- Come gestire i dati
 - analisi e logging centralizzato vs. distribuito

Esterno

- Rileva l'intero traffico diretto alla rete
- Più dati
- Più allarmi

Interno

- Rileva solo il traffico che entra effettivamente
- Verifica l'efficacia del firewall
- Non fornisce info sugli attacchi bloccati dal fw



Posizionamento nella rete aziendale



Sicurezza delle
reti

Monga

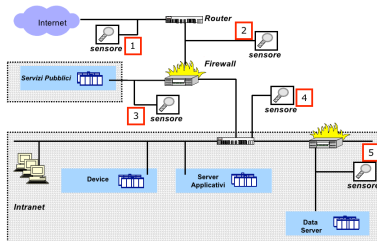
Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi
Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori



1. Esterno al border router

- Tutto il traffico diretto alla rete aziendale.
- Informazione completa e non filtrata.
- Tanti dati e allarmi

Posizionamento nella rete aziendale



Sicurezza delle
reti

Monga

Falsi allarmi

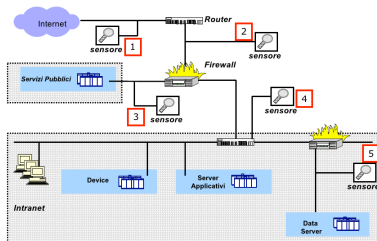
Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori



2. Tra border router e firewall

- Tutto il traffico meno quello filtrato
- Tanti dati, molti falsi allarmi.

Posizionamento nella rete aziendale



Sicurezza delle
reti

Monga

Falsi allarmi

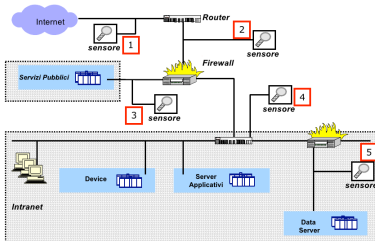
Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori



3. Sulla rete dei servizi pubblici, dietro il firewall

- Tutto il traffico autorizzato dal firewall e diretto ai servizi pubblici.
- Possibilità filtraggio mirato.
- Eventuale traffico illecito dai server pubblici.

Posizionamento nella rete aziendale



Sicurezza delle
reti

Monga

Falsi allarmi

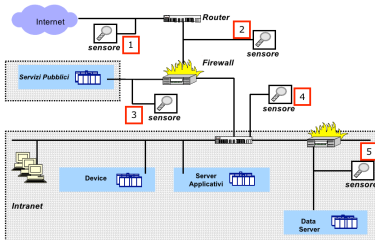
Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori



4. Sulla Intranet

- Sia il traffico da reti più esposte (es. DMZ) che interno alla Intranet.
- Rileva eventuali usi non leciti interni.
- Difficile dare firme, molti falsi allarmi.

Posizionamento nella rete aziendale



Sicurezza delle
reti

Monga

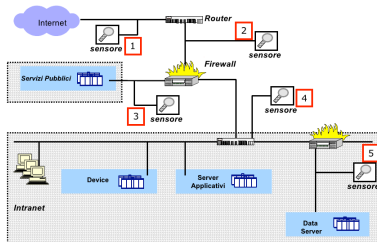
Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi
Teorema di
Bayes

Aspetti
architeturali

Posizionamento
sensori



5. Su di un segmento critico della rete aziendale

- Monitora le connessioni dirette ad alcune risorse particolarmente critiche della rete aziendale per le quali si richiede un livello di sicurezza più elevato (es. i server contenenti dati aziendali sensibili).
- Servizi specifici, quindi possibilità di configurazione mirata delle firme.



Sicurezza delle
reti

Monga

Falsi allarmi

Rilevamento
delle intrusioni

I falsi allarmi

Teorema di
Bayes

Aspetti
architettonici

Posizionamento
sensori

Il posizionamento dell'IDS influisce molto sulla sua efficacia

- ① Esterno
- ② Tra router e firewall
- ③ Vicino ai servizi
- ④ Sulla intranet
- ⑤ In punti critici