Rilevazione di anomalie di rete mediante analisi su serie temporali

Candidato:

Salvatore Costantino

Relatore:

Luca Deri

Introduzione

- Necessità di rilevare e mitigare le anomalie di rete
  - Aumento dei servizi offerti
  - Aumento delle problematiche legate alla sicurezza e alla gestione degli host connessi in rete
- Rilevazione di anomalie
  - Necessità di definire ciò che viene considerato normale
  - Registrazione di allarmi
- Mitigazione delle anomalie
  - Automatica
  - Manuale

Stato dell'Arte

- Signature-based IDS (Intrusion Detection System)

- Anomaly-based IDS
  - Statistical-based
  - Knowledge-based
  - Machine learning-based

- Rilevazione di anomalie su serie temporali
  - Analisi su un insieme di serie temporali
  - Analisi su singola serie temporale

Obiettivi

- Realizzazione di un sistema automatico di rilevazione di anomalie su serie temporali
- Implementazione tecnica di mitigazione
- Efficiente in spazio ed in tempo, in modo da poter analizzare gli host di un’intera sottorete
- Apprendimento del comportamento passato, per le metriche a lungo termine
- Confronto con un profilo comportamentale considerato normale, per le metriche a breve termine
- Buona performane in termini di precisione, specificità e sensibilità

Motivazione

- Attualmente alcune tecniche risultano troppo costose per analizzare un’intera sottorete (reti neurali), specialmente nel breve termine

- Effettuare ragionamenti preliminari sui dati da analizzare, fase trascurata da molti articoli individuati in letteratura

- Enfasi posta sull’analisi di serie temporali, in modo da studiare il comportamento temporale di una metrica di rete. Nella letteratura il fattore tempo viene spesso tralasciato.

Contributo originale

- Uso di algoritmi e metriche non attualmente presenti nello stato dell'arte relativo alla rilevazione di anomalie di rete:
  - **Indicatore statistico finanziario**, nel breve termine
  - **Modello di predizione** su serie temporale, nel medio-lungo termine
  - **Utilizzo e correlazione di un ampio insieme di metriche**, grazie al software di monitoraggio di rete ntop

Metriche

- Metriche a breve termine (analizzate dal sistema ogni 5 minuti), ovvero coppie di metriche il cui rapporto in situazioni normali si mantiene più o meno costante nel tempo e non supera alcuni valori soglia

- Metriche a medio-lungo termine (analizzate dal sistema ogni ora), ovvero singola metrica avente solitamente trend e stagionalità

- Le metriche considerate compaiono molto spesso nel traffico odierno e si è visto che esse sono spesso affette da anomalie

Metriche a Breve Termine

Data la coppia di contatori (x, y) relativi alle metriche (M_x, M_y) da analizzare, consideriamo

$$\frac{\Delta x}{\Delta y} = \frac{x_{t_f} - x_{t_i}}{y_{t_f} - y_{t_i}} = r_{t_i}, \quad \Delta r = r_{t_f} - r_{t_i}$$

- In totale sono state analizzate 14 metriche a breve termine tra cui:
  - $\frac{risposte\ DNS\ ricevute}{richieste\ DNS\ inviate + risposte\ DNS\ ricevute}$
  - $\frac{bytes\ protocollo\ DNS\ ricevuti}{pacchetti\ DNS\ ricevuti}$
  - $\frac{flussi\ sospetti\ come\ client}{flussi\ totali\ come\ client}$

Metriche a Medio-lungo Termine

- Singole metriche aventi **trend** e **multi-stagionalità**

- Bytes inviati, Bytes ricevuti, flussi come client, flussi come server

- Ulteriore controllo sulle seguenti categorie di protocolli/eventi:
  - Accesso Remoto
  - Protocolli Sconosciuti
  - Malware
  - Mining

Rilevazione delle Anomalie

- Rilevare un’anomalia significa individuare eventi o valori che per qualche loro caratteristica non possono essere considerati normali

- Criteri di normalità per le metriche a breve termine:
  - Non superamento dei valori soglia
  - Comportamento (valore delle metriche) più o meno costante nel tempo

- Criterio di normalità per le metriche a breve termine:
  - Comportamento futuro coerente con quello passato

Threshold

- Tecnica applicata a tutte le metriche a breve termine

- Soglia fissata a **576** (bytes) per le metriche relative alla dimensione media dei pacchetti DNS inviati e ricevuti
  - Data exfiltration/infiltration

- Soglia fissata a **0.50** per le altre metriche a breve termine
  - Valore rapporto anomalo

RSI

- Indicatore statistico utilizzato per effettuare analisi su mercati finanziari, in grado di rilevare la velocità del movimento dei prezzi
- Idea chiave: utilizziamolo per misurare la velocità con cui variano i valori legati alle metriche di rete
- Tecnica applicata a quasi tutte le metriche a breve termine
  - Dopo aver applicato la tecnica delle soglie fisse

- $RSI = 100 * U/(U + D)$
  - $U = \frac{\sum_{t=2}^{N+1} \max(0, x_t - x_{t-1})}{N}, D = \frac{\sum_{t=2}^{N+1} |\min(0, x_t - x_{t-1})|}{N}$
  - $U = (U_{old} * (N - 1) + \max(0, x_{new} - x_{last}))/N$
  - $D = (D_{old} * (N - 1) + |\min(0, x_{new} - x_{last})|)/N$
- Oscilla tra due valori: 0 e 100

Threshold & RSI

- Condizioni volumetriche minime di traffico
  - Valori minimi delle metriche, affinché i valori registrati risultino significativi

- Condizioni su traffico p2p
  - Non vengono effettuate analisi sui messaggi ICMP port unreachable e host unreachable, se è presente traffico p2p (in tal caso i valori risultano fisiologicamente alterati)

Prophet

- Modello di regressione che può assumere una delle seguenti forme:
  - $y(t) = g(t) + s(t) + h(t) + \varepsilon_t \ (additivo)$
  - $y(t) = g(t) * s(t) + h(t) + \varepsilon_t \ (moltiplicativo)$
- g lineare, $s(t) = \sum_{n=1}^{N} (a_n \cos(2\pi nt/P) + b_n \sin(2\pi nt/P))$
- Scelta iper-parametri
  - changepoint_prior_scale
  - seasonality_prior_scale
  - ordine serie di Fourier
- Usiamo il modello moltiplicativo
- Tempo per training e predizione: circa 9 secondi
- Intervalli di incertezza integrati, per la rilevazione di anomalie

Allarmi

TYPE	ANOMALY	HOST/MAC	IF	DATE	METHOD	VAL
START	ping_packets	[redacted].239@125	0	2019-06-27T20:15:00Z	TRESHOLD	1.0
START	ping_packets	[redacted].145@125	0	2019-06-27T10:35:00Z	TRESHOLD	0.5
END	ping_packets	[redacted].145@125	0	2019-06-27T17:00:00Z	TRESHOLD	
START	ping_packets	[redacted].180@125	0	2019-06-27T13:40:00Z	TRESHOLD	0.6
END	ping_packets	[redacted].180@125	0	2019-06-27T13:45:00Z	TRESHOLD	
START	ping_packets	[redacted].241@125	0	2019-06-27T09:15:00Z	TRESHOLD	0.5
END	ping_packets	[redacted].241@125	0	2019-06-27T10:00:00Z	TRESHOLD	
START	ping_packets	[redacted].241@125	0	2019-06-27T10:20:00Z	TRESHOLD	0.6
END	ping_packets	[redacted].241@125	0	2019-06-27T10:25:00Z	TRESHOLD	
START	ping_packets	[redacted].241@125	0	2019-06-27T12:35:00Z	TRESHOLD	0.5
END	ping_packets	[redacted].241@125	0	2019-06-27T12:55:00Z	TRESHOLD	
START	ping_packets	[redacted].78@125	0	2019-06-27T08:15:00Z	TRESHOLD	0.6
START	ping_packets	[redacted].234@125	0	2019-06-27T09:25:00Z	TRESHOLD	0.8
END	ping_packets	[redacted].234@125	0	2019-06-27T21:25:00Z	TRESHOLD	
START	ping_packets	[redacted].234@125	0	2019-06-27T21:40:00Z	TRESHOLD	0.5

Mitigazione tramite XDP

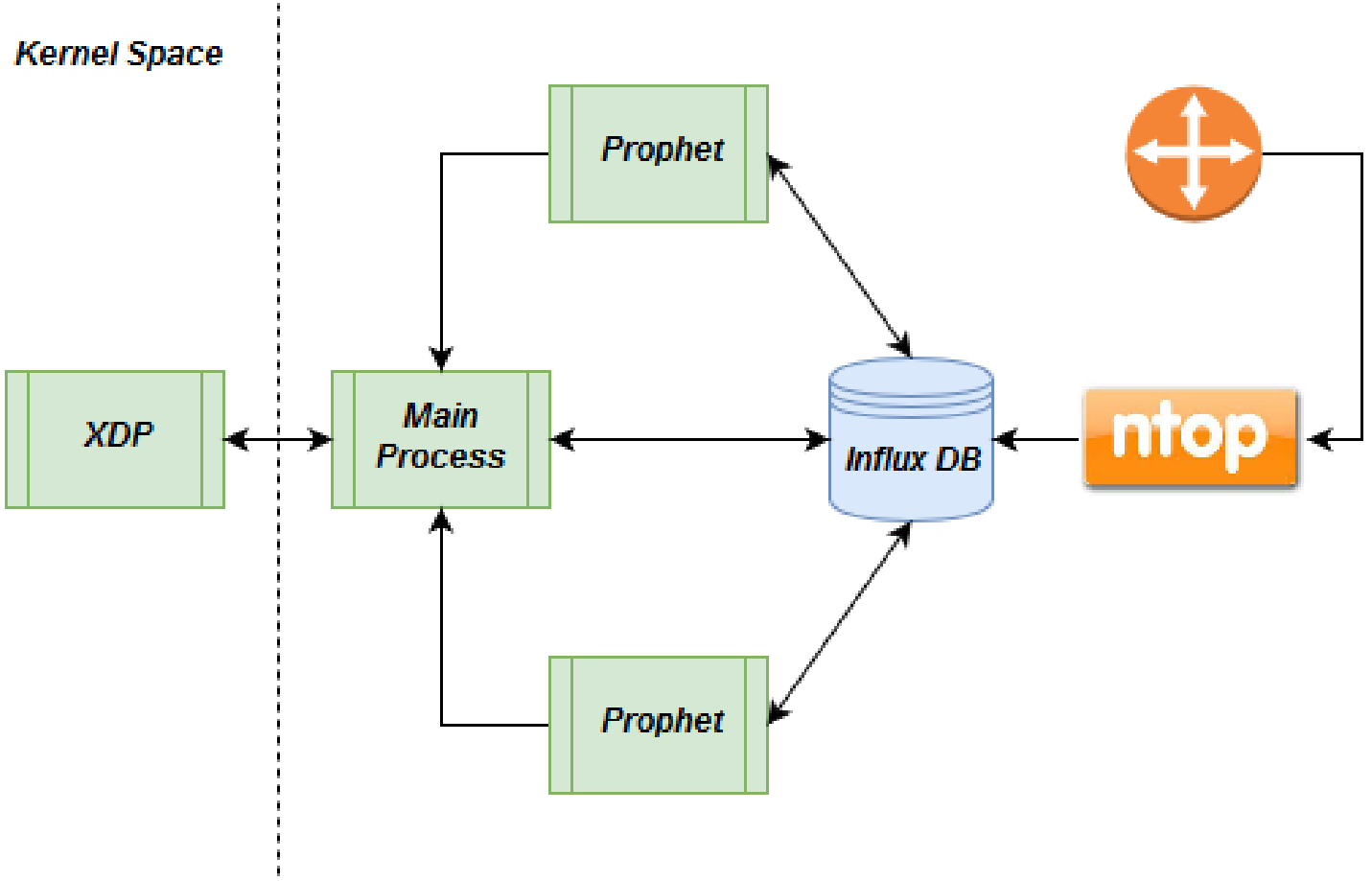
- Analisi dei pacchetti direttamente all'interno del Kernel Linux, grazie alla tecnologia **eBPF** (Extended Berkeley Packet Filter)
- Eseguito nella parte bassa dello stack protocollare
  - Velocità di filtraggio molto elevata
- Il codice XDP, prima di essere iniettato nel Kernel Linux, deve essere validato da un verificatore:
  - Non sono ammessi cicli
  - controllo esplicito dei limiti di memoria del pacchetto sotto analisi

Controllo limiti memoria

```
void* data_end = (void*)(long)ctx->data_end;
void* data = (void*)(long)ctx->data;

struct ethhdr *eth = data; //struct header ethernet
uint64_t nh_off = sizeof(*eth);
if (data + nh_off > data_end) return XDP_DROP; //check bounds
uint64_t macIn = mac2u64(eth->h_source);
uint64_t macEg = mac2u64(eth->h_dest);
if(checkMac(&macIn) || checkMac(&macEg)) return XDP_DROP;
uint16_t h_proto = eth->h_proto;
```

Architettura software



Validazione dei modelli

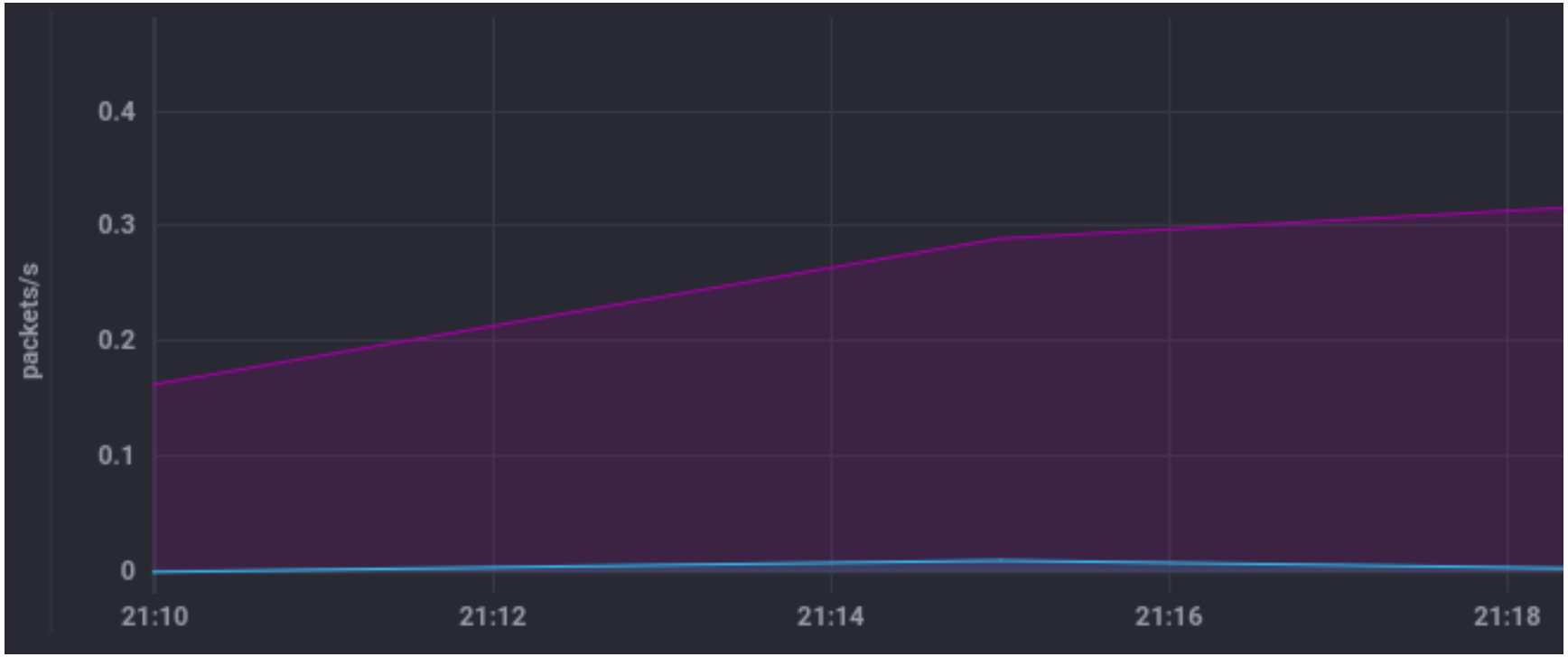
- Valori Soglia
  - Scelto valore soglia a 576 per metriche relative alla dimensione media dei pacchetti DNS
  - Scelto valore soglia a 0.50 per le altre metriche

- Periodo e valore soglia RSI
  - Scelto periodo pari 50
  - Scelta soglia pari a 80

- Prophet
  - Model Selection per scelta iper-parametri
  - Scelta del modello moltiplicativo

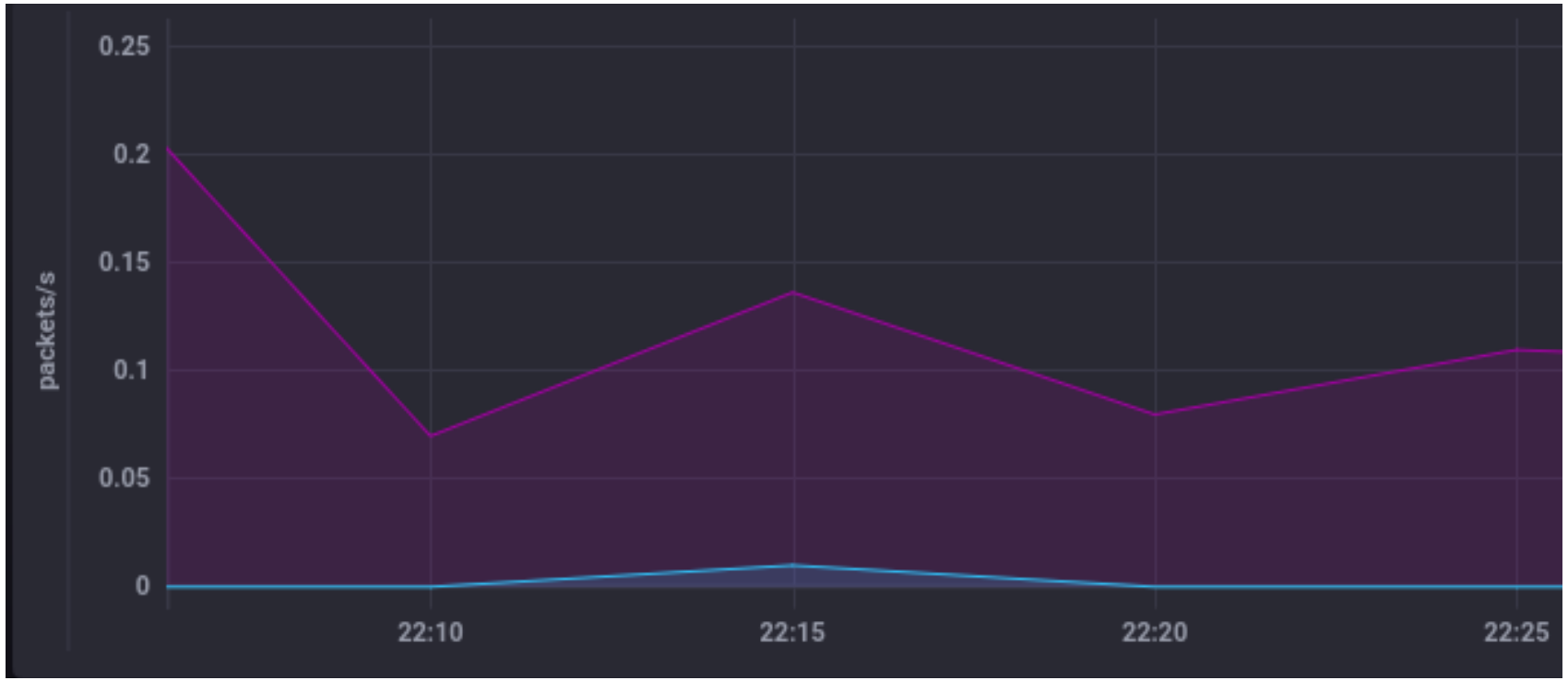
Periodo RSI: 25 vs 50

- Più è corto il periodo, più l'RSI risulta sensibile con il rischio di generare falsi allarmi (falsi positivi)
- Falso positivo rilevato di tipo «dns_errors» (rapporto circa 0.03)

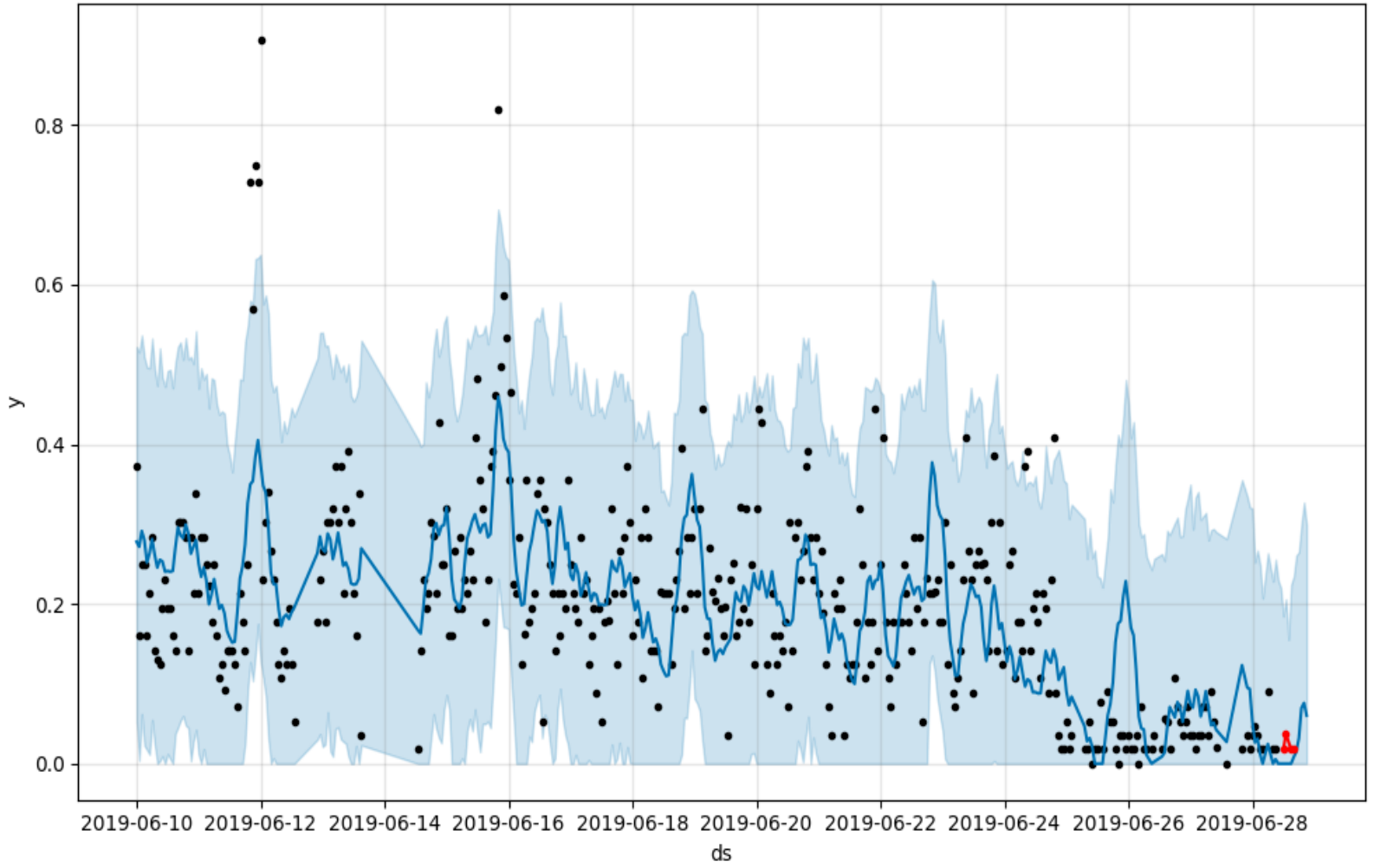


Soglia RSI: 80 vs 70

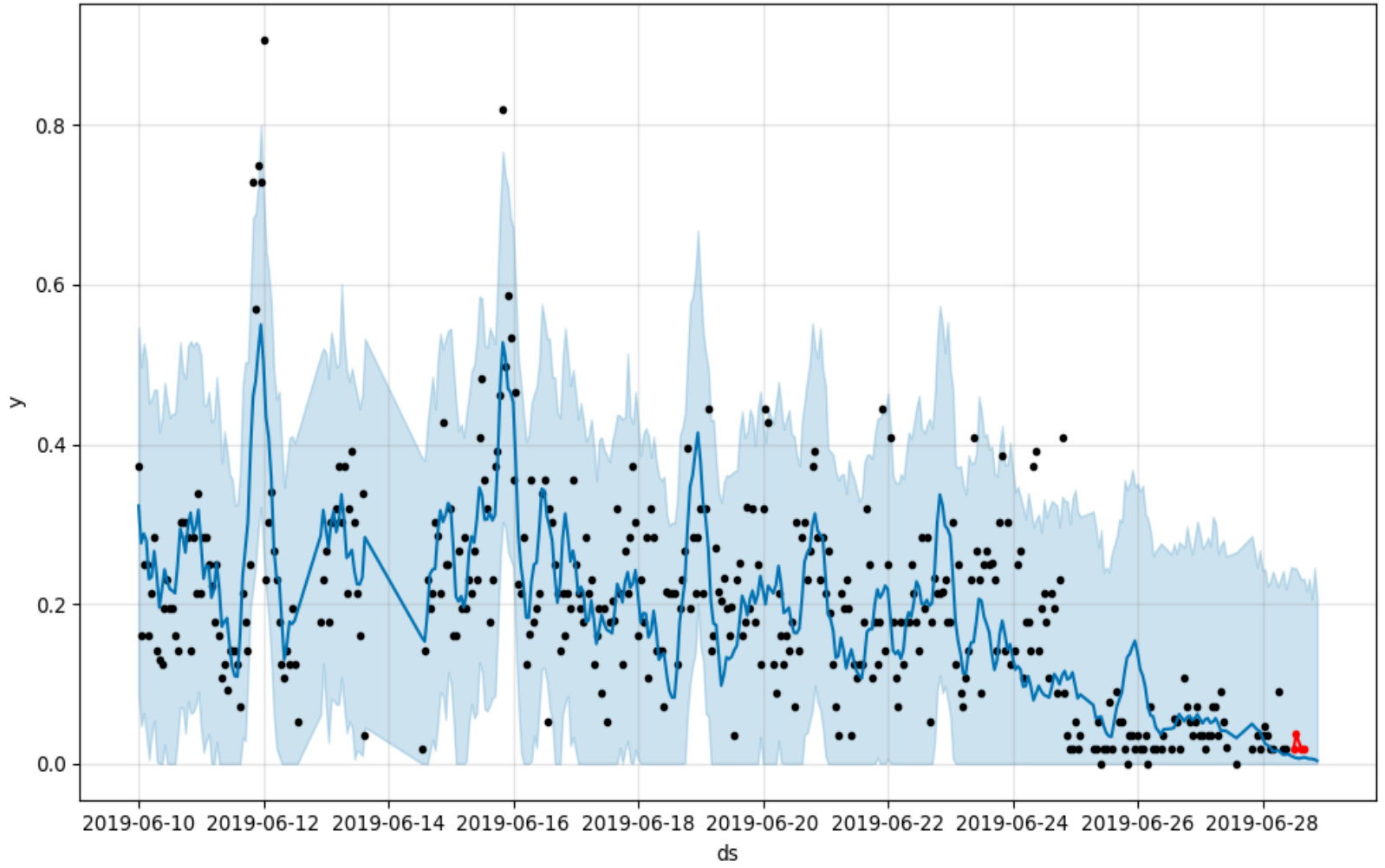
- 70 è la soglia superiore consigliata, ma essa genera vari falsi positivi
- Falso positivo rilevato di tipo «dns_errors» (rapporto circa 0.07)



Modello Prophet: Additivo



Modello Prophet: Moltiplicativo



Validazione Performance di Rilevazione

- Validazione effettuata sugli host di un ISP locale, contenente decine di migliaia di host

- Si sono considerati alcuni host campione, in base ai risultati del report di rilevazione di anomalie

Statistiche generali

TYPE	TOTAL_CHECK	ANOMALIES	METHOD
ping_packets	2602	1684	TRESHOLD
dns_packets	35406	33420	TRESHOLD
dns_errors	35299	4586	TRESHOLD
port_unreach_srv	4237	33	TRESHOLD
port_unreach_clt	9355	8	TRESHOLD
host_unreach_clt	9054	480	TRESHOLD
host_unreach_srv	1891	0	TRESHOLD
TCP_client_iss	1133	377	TRESHOLD
TCP_server_iss	1119	19	TRESHOLD
dns_size_srv	474	0	TRESHOLD
dns_size_clt	444	0	TRESHOLD
anmls_flows_srv	29644	949	TRESHOLD
anmls_flows_clt	33714	2040	TRESHOLD
dns_errors	14666	15	RSI
port_unreach_srv	2555	9	RSI
port_unreach_clt	5285	13	RSI
host_unreach_clt	4273	12	RSI
host_unreach_srv	1191	4	RSI
TCP_client_iss	186	0	RSI
TCP_server_iss	573	0	RSI
dns_size_srv	295	0	RSI
dns_size_clt	265	0	RSI
anmls_flows_srv	13461	0	RSI
anmls_flows_clt	14844	0	RSI
flows_as_client	28	2	PROPHET
flows_as_server	199	0	PROPHET
bytes_sent	28	0	PROPHET
bytes_rcvd	199	1	PROPHET

Statistiche per host

HOST	TYPE	TOTAL_CHECK	ANOMALIES	METHOD
[redacted].130@125				
	dns_errors	97	0	RSI
	anmls_flows_srv	59	0	RSI
	anmls_flows_clt	100	0	RSI
	dns_packets	148	59	TRESHOLD
	dns_errors	148	1	TRESHOLD
	port_unreach_clt	1	0	TRESHOLD
	TCP_client_iss	25	4	TRESHOLD
	TCP_server_iss	27	0	TRESHOLD
	anmls_flows_srv	109	0	TRESHOLD
	anmls_flows_clt	150	0	TRESHOLD

Validazione Tecnica Threshold

- Controllati 25 host: 15 host risultati anomali e 10 host risultati non anomali

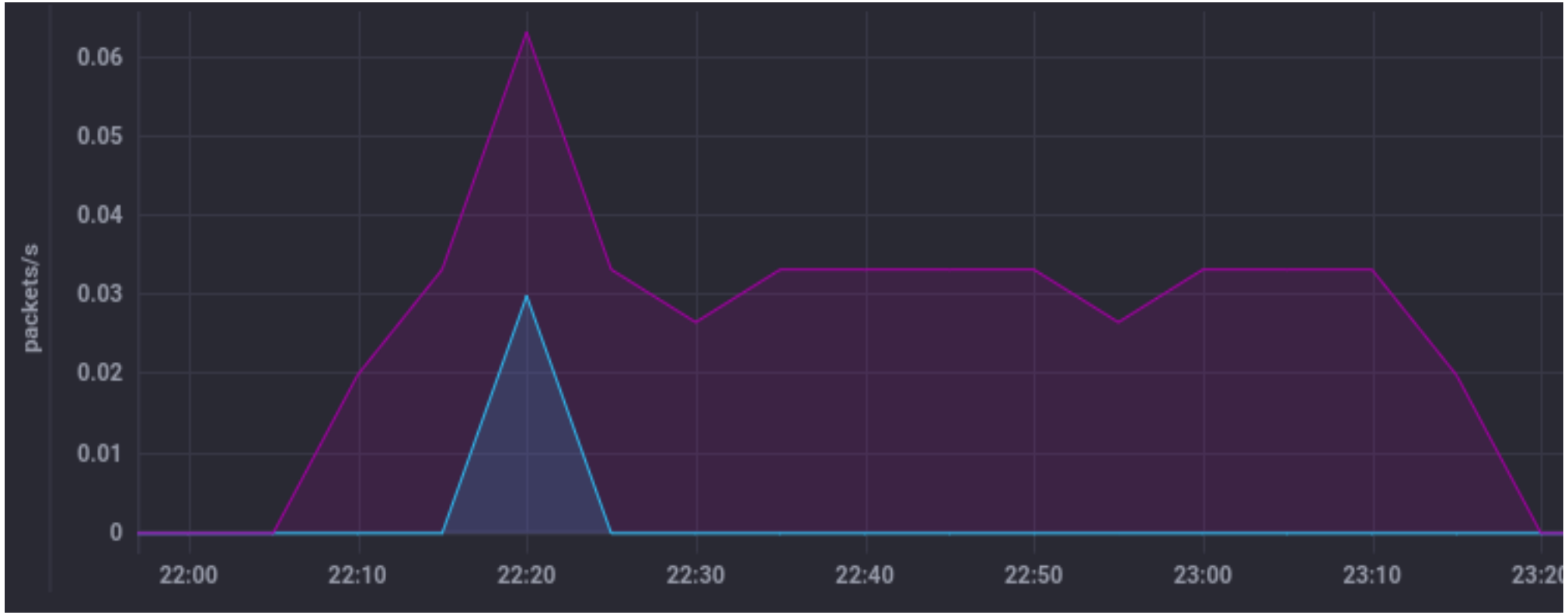
$$PRECISIONE = \frac{TP}{TP + FP} = \frac{13}{13 + 2} = 87\%$$

$$SPECIFICITA' = \frac{TN}{TN + FP} = \frac{10}{10 + 2} = 83\%$$

$$SENSIBILITA' = \frac{TP}{TP + FN} = \frac{13}{13 + 0} = 100\%$$

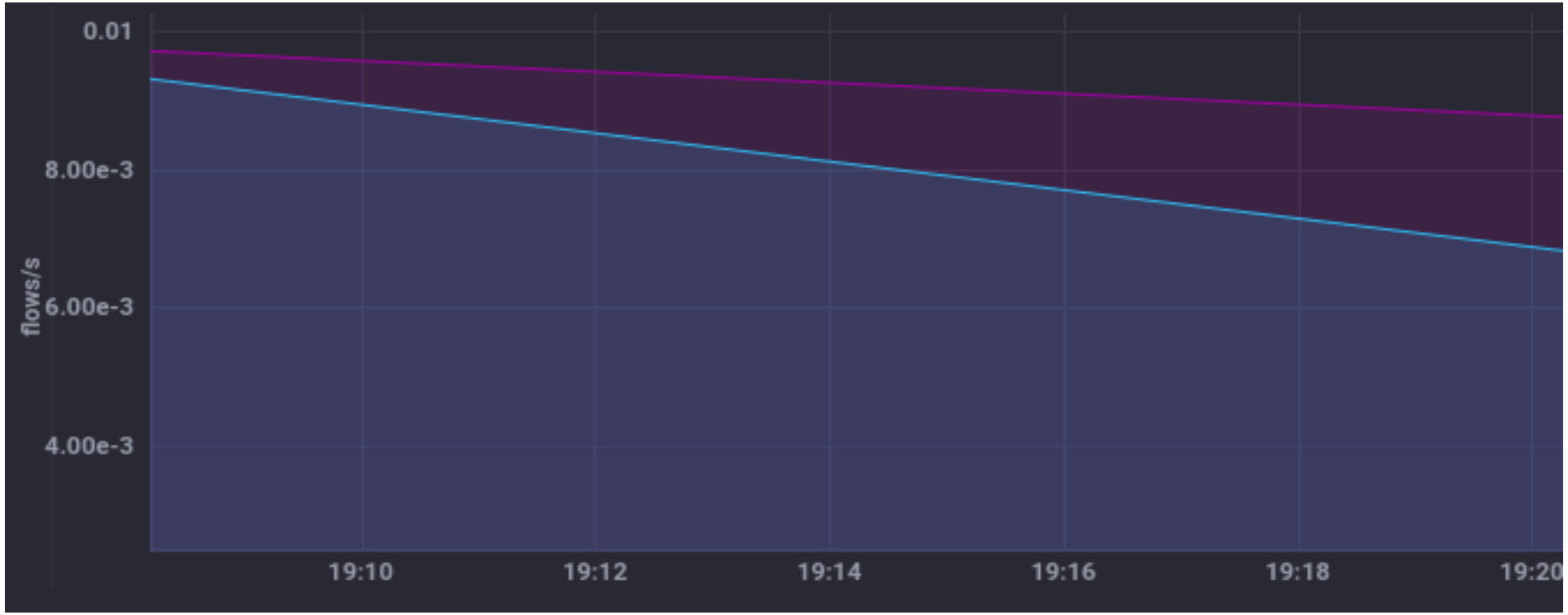
Vero Positivo Threshold

- «ping_packets»: presenti risposte (curva viola), senza richieste (curva blu)



Falso Positivo Threshold

- «anomalous_flows_as_client»: flussi sospetti come client (curva blu) < 0.01 (flussi al secondo)



Validazione RSI

- Controllati 20 host: 10 host rilevati anomali, 10 host non rilevati anomali

	Host anomali	Host non anomali
Rilevati	7	3
Non rilevati	2	8

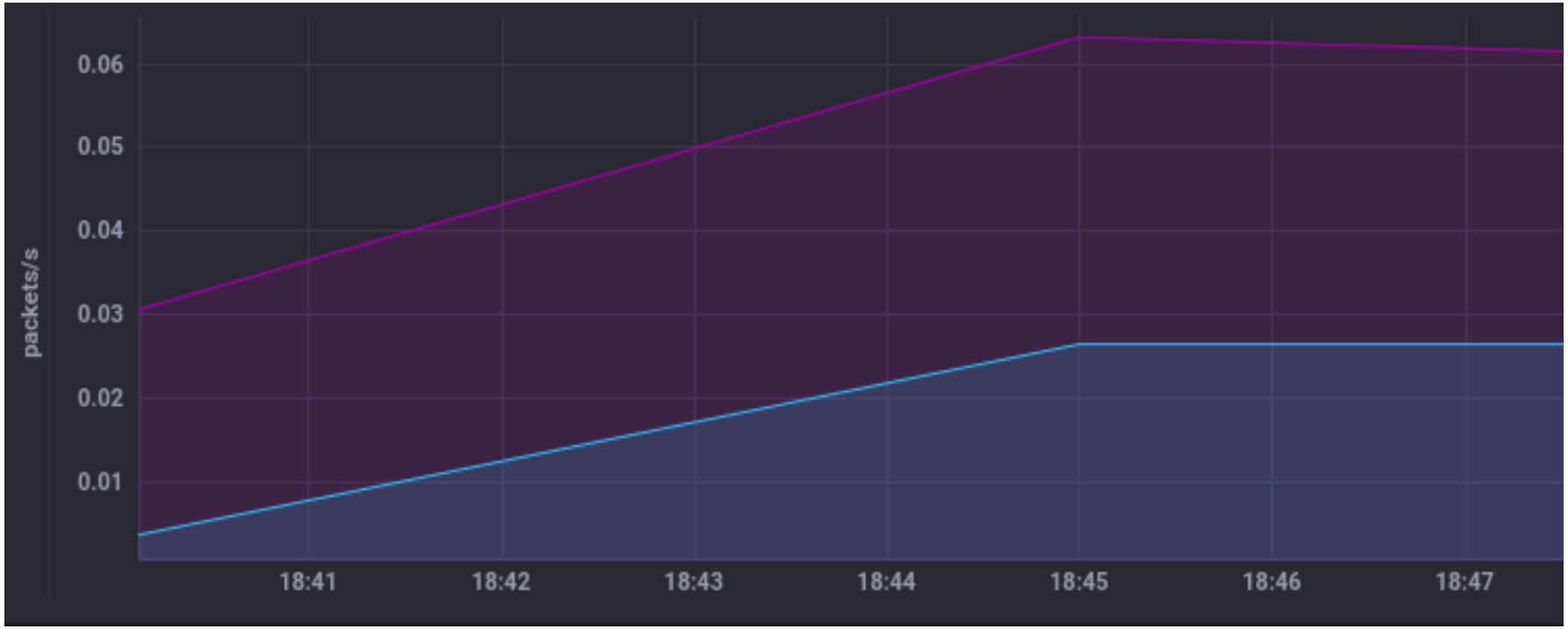
$$PRECISIONE = \frac{TP}{TP + FP} = \frac{7}{7 + 3} = 70\%$$

$$SPECIFICITA' = \frac{TN}{TN + FP} = \frac{8}{8 + 3} = 73\%$$

$$SENSIBILITA' = \frac{TP}{TP + FN} = \frac{7}{7 + 2} = 78\%$$

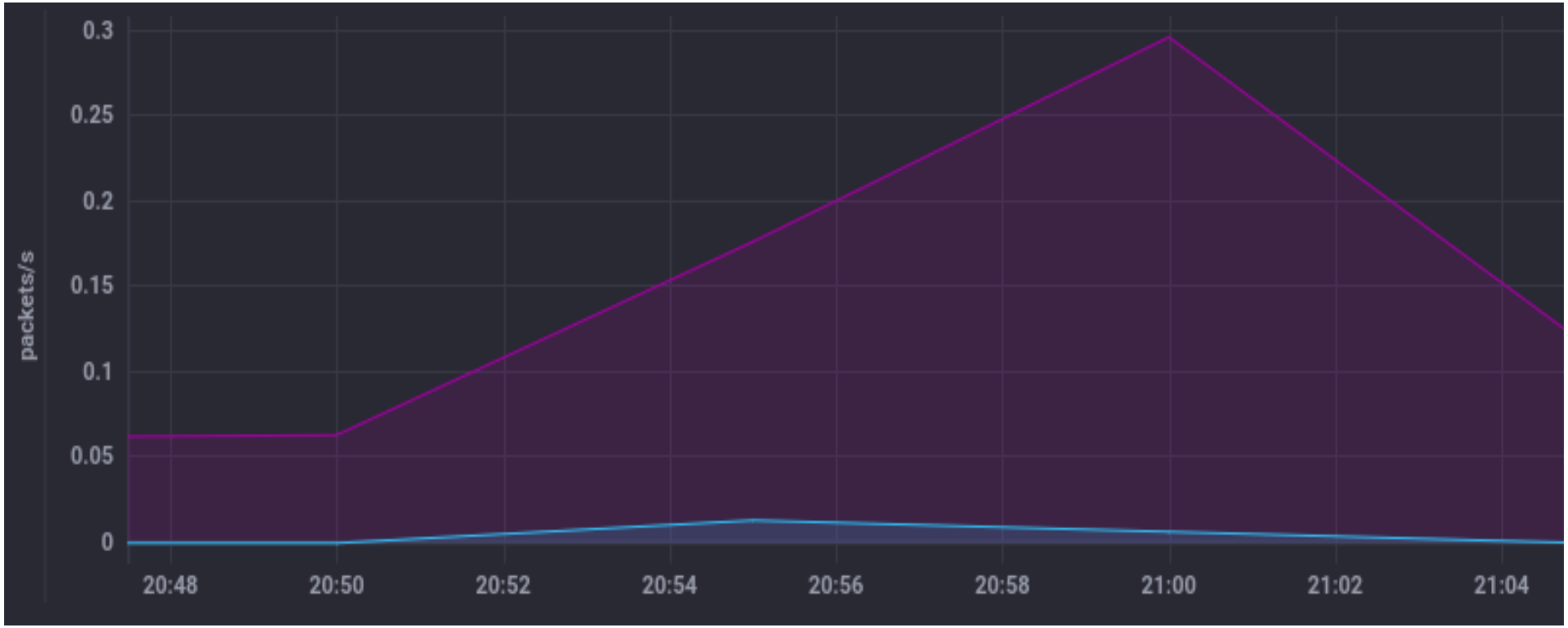
Vero Positivo RSI

- «dns_errors»: si passa da 0 a circa 0.50 (valore rapporto)



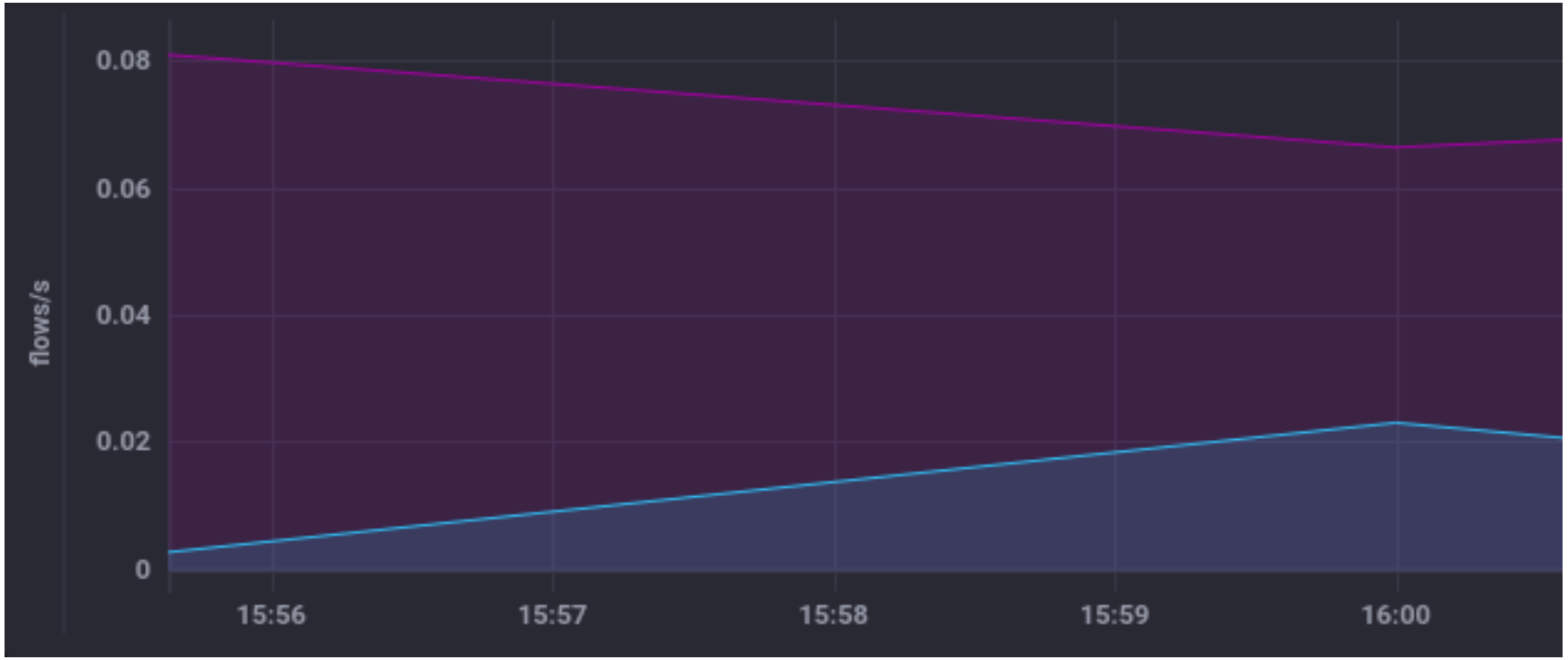
Falso Positivo RSI

- «dns_errors»: rapporto non significativo (< 0.15)



Falso Negativo RSI

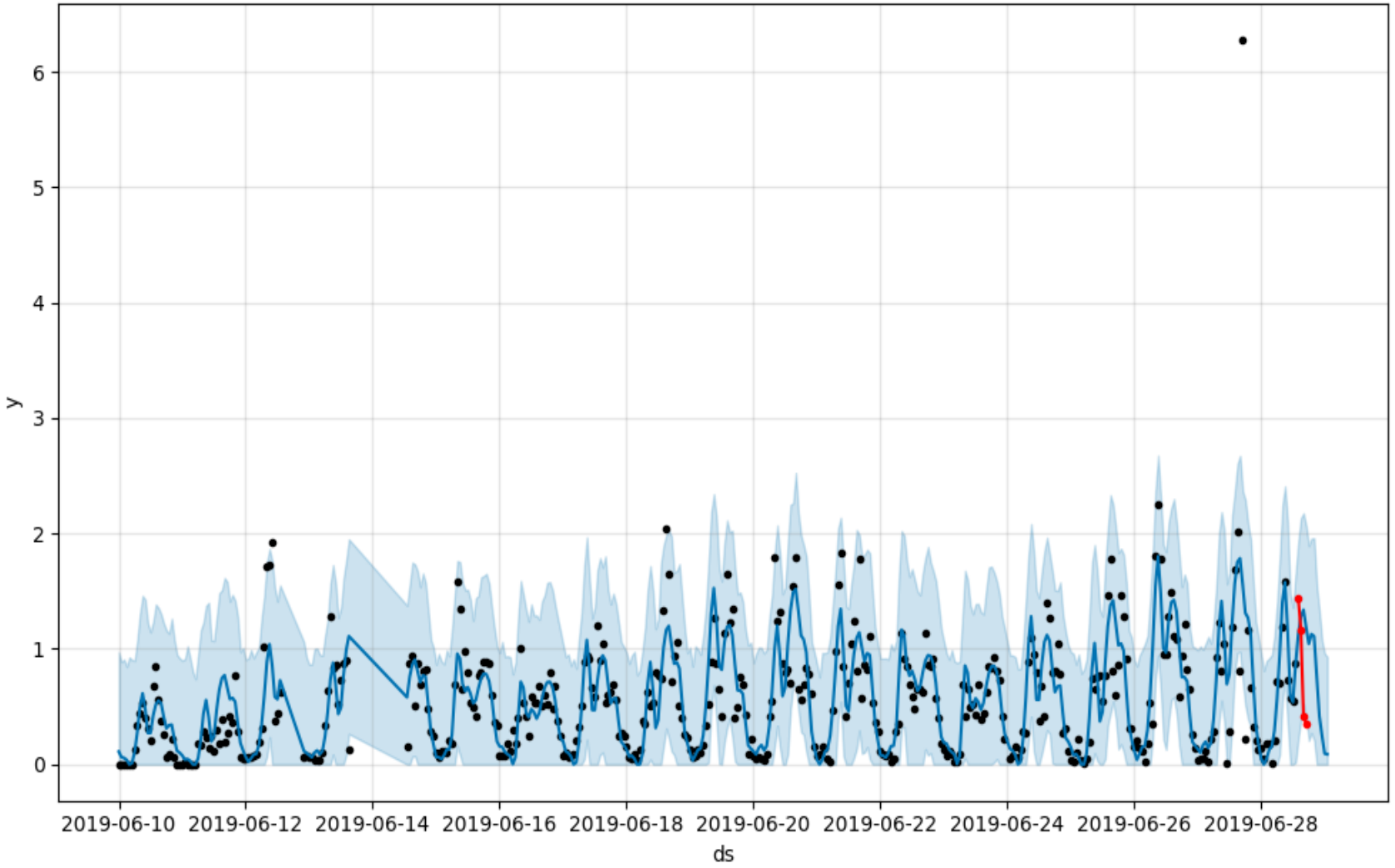
- «host_unreachable_as_client»: anomalia presente all'interno dei primi 51 valori, su cui calcolare l'RSI (il trend generale risulta costante)



Validazione Prophet

- 3 anomalie rilevate, su oltre 400 controlli
- Tutte e 3 sono falsi positivi, riconducibili ad un cambiamento di comportamento fisiologico degli host analizzati
- I 3 falsi positivo non vengono rilevati, se viene attivato il controllo delle categorie NDPI

Falso positivo Prophet



Risultati finali

		Host anomali	Host non anomali
Threshold	Rilevati	**13**	**2**
	Non rilevati	**0**	**10**
RSI	Rilevati	**7**	**3**
	Non rilevati	**2**	**8**
Prophet + NDPI	Rilevati	**0**	**0**
	Non rilevati	**0**	**454**
Totale	Rilevati	**20**	**5**
	Non rilevati	**2**	**472**

$$PRECISIONE = \frac{TP}{TP + FP} = \frac{20}{20 + 5} = 80\%$$

$$SPECIFICITA' = \frac{TN}{TN + FP} = \frac{472}{472 + 5} = 99\%$$

$$SENSIBILITA' = \frac{TP}{TP + FN} = \frac{20}{20 + 2} = 91\%$$

Lavori Futuri

- Correlazione tra serie temporali di host diversi
- Allarmi come input ad un livello di analisi superiore, per esempio un autoencoder
- Miglioramento tecnica di mitigazione, che appare troppo drastica e non in grado di proteggere un eventuale host sotto attacco

Conclusione

- Il problema della rilevazione di anomalie non ha, ad oggi, una soluzione semplice e universale, e ogni tecnica presenta i suoi punti di forza e debolezza
- Si è realizzato un sistema intelligente, capace di analizzare e mitigare alcune anomalie di rete presenti in un'intera rete, in modo efficiente e con buona precisione
- Si noti, che in questo lavoro di Tesi vengono analizzati host generici, e non è possibile effettuare alcuna assunzione sul tipo di traffico che essi possono generare (come può accadere in ambiente IoT (Internet of Things)); quindi il problema che si è trattato è risultato particolarmente difficile da affrontare