



Difesa

Anomaly Detection Systems

Fabio Mostarda Pierpaolo Palazzoli

Grado di difficoltà



XXXXXXXXXXXXXXXX

Nei processi di messa in sicurezza di reti complesse ci si trova a dover approcciarsi con realtà complesse e molto varigate, la presenza ormai sul mercato di applicazioni e sistemi sempre più complessi nel codice e nell'implementazione hanno messo in seria difficoltà i security manager IT. La stratificazione dei servizi applicativi di sicuro non aiuta nelle policy di sicurezza rendendo il mondo della comunicazione telematica un mare di caos caratterizzato da andamenti di servizio altalenanti. Sistemi di sicurezza proattivi aiutano nella gestione ma si legano a implementazioni di regole: le regole come nella vita, vengono scritte ed aggiornate ma sono statiche e quindi vivono il loro ciclo funzionale solo nel momento nel quale vengono applicate. Da qui viene l'esigenza di introdurre metodi di approccio all'analisi di attacchi e intrusioni, tramite mezzi statistici in grado di rilevare delle discrepanze. Il nome che contraddistingue questa analisi è Anomaly Detection. L'approccio a questa tecnica, molto discussa, passa da modelli matematici statistici o al limite da metodi di analisi. Lo strumento in grado di aiutare maggiormente colui che gestisce la security è un programma in grado di esportare in maniera

più semplice possibile possibile lo "stato d'arte" della propria infrastruttura storicizzandone i protocolli in gioco. Non esistono sistemi proattivi di anomaly detection, cioè sistemi in grado, senza interfacciamento con l'operatore, di rilevare attacchi: l'intervento umano in questo tipo di attività risulta fondamentale, in quanto variazioni statistiche tipicamente sono caratterizzate da implementazioni di nuovi servizi che non necessariamente vengono erogati dalla stessa infrastruttura. La natura interconnessa della rete permette a utenti di potersi fornire di servizi che non sono necessariamente appartenenti al proprio gestore di connettività, si aggiunge poi il fatto che gli attacchi vengono tipicamente scritti da persone e quindi persona-

Dall'articolo imparerai...

- conoscere l'importanza e l'utilità delle tecnologie di Anomaly Detection.
- imparare a proteggersi da potenziali aggressori.

Cosa dovresti sapere...

- Minima conoscenza di linux.

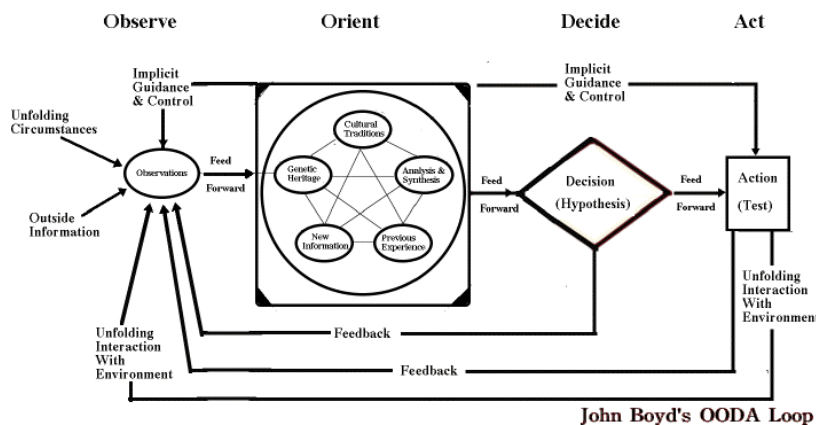


Figura 1. XXXXXXXXXXXX [Fig 1 OODA.gif]

lizzabili anche nelle tempistiche e nel mezzo. E' certo che la maggior parte di questi sono eseguiti in modalità automatizzata, ma come ci insegna il mondo dei virus, il poliformismo è un tipico mezzo finalizzato a confondere le analisi.

Per poter comprendere appieno il mondo di cui si parla bisogna avere una idea di massima dei protocolli che si ha la certezza di gestire sulla propria infrastruttura, il significato funzionale dei propri server e le possibilità, intese come libertà di azione, date ai propri client.

L'importanza del monitoraggio come base del processo di reazione

In molti contesti si ritiene sufficiente, per soddisfare le necessità di sicurezza di un'organizzazione, predisporre un'infrastruttura di protezione basata su firewall, sopra al quale vengono introdotti meccanismi di hardening software o di intrusion prevention (IPS).

Avere un tale sistema stratificato di difesa è senza dubbio un requisito necessario nello scenario attuale, ma non è detto che sia sempre sufficiente; come per ogni sistema ideato dall'uomo, anch'esso può fallire: su questo non vi sono dubbi. Le vere domande da porsi sono: come accorgersi tempestivamente del fallimento, come rispondere nel modo migliore alla minaccia; la vera sfida è riuscire a evitare o limitare i danni, nonostante un bypass parziale delle proprie difese.

Per quanto difficile, questa impresa non è persa in partenza.

A livello teorico, una prima risposta potrebbe essere quella di spingere per una modellazione sempre più precisa della realtà operativa (vulnerabilità software, minacce, attacchi), al fine di costruire un modello concettuale il più realistico e predittivo possibile; questo sarebbe poi implementabile in un "super software" in grado di rendere invulnerabile l'infrastruttura da difendere.

Questo approccio, che potrebbe sembrare il migliore e il più risolutivo, rischia di divenire fallimentare se non tenesse in considerazione un dettaglio: la gestione delle anomalie e delle eccezioni. Si potrebbe erroneamente pensare che, in quanto tali, siano di poco conto, o addirittura eliminabili con una maggior precisione nel modello.

Purtroppo questo non è esatto: come dimostra Goedel, ogni sistema coerente è incompleto, e d'altro canto la coerenza di un sistema deduttivo non è dimostrabile; come aveva notato John Boyd, questo è un aspetto della realtà che è molto simile al secondo principio della termodinamica: in un sistema chiuso, l'entropia non può diminuire. Il caos, le anomalie, esistono e non sono eliminabili se non si rende il sistema disponibile ad accettare feedback esterni; la rilevazione delle anomalie diventa quindi molto importante.

Proviamo per un istante ad applicare, nel "processo decisionale" della sicurezza informatica, un modello strutturato in grado di contemplare

il caos e l'incertezza; proposto da John Boyd negli anni '70, è noto nel mondo militare come "OODA-Loop". (vedi figura1).

Esso si basa sui seguenti step:

Observation (osservazione dei fenomeni e delle anomalie)

Orientation (la fase più delicata, ovvero la sintesi delle informazioni provenienti dall'osservazione)

Decision

Action

Come si può notare, le decisioni e le azioni da intraprendere sono subordinate all'orientamento preso; quest'ultimo è influenzato dalle skill e dal background degli operatori, che però a loro volta devono basare le proprie valutazioni su ciò che osservano. L'osservazione è anche la base di tutti i feedback che rendono il modello flessibile e non statico e chiuso.

Aumentare il grado di conoscenza e consapevolezza dello stato della propria infrastruttura informatica è fondamentale per permettere una risposta strutturata a problemi nuovi; questo significa, scendendo più in dettaglio, avere un'idea di cosa si debba intendere come "funzionamento normale" (della propria rete e dei propri server/servizi) e, di conseguenza, di cosa sia invece "anomalo".

Questa attività deve permettere la generazione di "pattern standard di comportamento", suddivisi nel modo più opportuno (di traffico, di protocollo, su base oraria, per servizio, etc..), nonché delle "blacklist" di elementi sospetti o che dovrebbero essere del tutto assenti.

Anomaly Detection: gli strumenti

Nel vasto panorama del software per il monitoraggio delle reti, concentreremo la nostra attenzione su due prodotti gratuiti ed open source che, grazie alla qualità dei risultati e alla loro flessibilità, hanno goduto di una grande diffusione e popolarità. Questi software sono NTOP e OURMON.



Global TCP/UDP Protocol Distribution

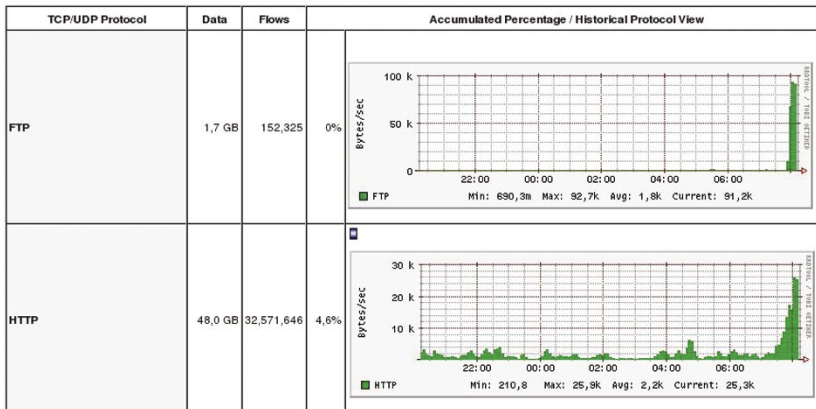


Figura 3. XXXXXXXXXXXXXXXXXXXX [Fig 3. TCP_UDP.jpg]

Ntop

Ntop è un uno strumento di monitoraggio in grado di dettagliare l'attività protocollare. Le sue funzionalità garantiscono una impronta dettagliata della rete dal livello 2 al livello 7 ISO/OSI. L'interfaccia utente è web, rendendolo estremamente portabile. Il software analizza il traffico tramite le librerie libpcap, e in alternativa tramite esportazioni statistiche NetFlow e SFlow. Integra implementazione di grafici tramite le rrdtool garantendo un effetto vistoso di variazioni di traffico.

Il pacchetto è ormai integrato e in qualsiasi distribuzione, i formati disponibili sono .deb, .rpm e sorgente. In fase di installazione l'unica attività da compiere da linea di comando è

semplicemente lanciare il programma la prima volta per dare la password di admin, questo utente protegge la parte di configurazione, del pacchetto stesso, con una password.

Lanciato con il proprio init script il programma di default apre un proprio web server sulla porta tcp 3000; collegandosi con un browser si otterrà l'interfaccia in (vedi figura 2).

La pagina introduttiva organizza i dati di analisi prelevati con analisi di pacchetto (libpcap) o di flusso (Sflow e NetFlow). Le statistiche di sniffing sono tipicamente contraddistinte da un contatore legato all'interfaccia che si intende usare in sniffing, nel caso si usi una tipologia di dati di flow comparirà un'interfaccia virtuale.

I dati più significativi raccolti da questo strumento sono di si-

curo quelli legati alla tipologia di traffico.(Fig Protocolli.jpg) Nel grafico vengono messi graficamente alcune tipologie di traffico applicativo di modo da far denotare quali siano le applicazioni più utilizzate. La discrepanza rilevabile, con una giusta storizzazione, è un picco di qualche applicazione specifica. La legenda del traffico tiene sempre segnati i valori di minimo, massimo e current, questi potrebbero segnalare qualche strana anomalia. I grafici visualizzati sono tutti creati tramite rrdtool che è in grado di avere buona accuratezza nei valori.

Le sezioni di enumerazione delle connessioni TCP/UDP sono molto dettagliate(vedi figura 3)

di modo da poter sapere le percentuali di tipologia di connessioni, notando così cambiamenti repentini e non giustificabili da nessun evento o nuova attivazione.

Le sezioni tipiche di un sistema di monitoraggio quali banda passante e carico di rete (Fig Load.jpg) sono storizzate di modo da avere presente i trend di passaggio anche mensili.

Le analisi statistiche per host e per direzione del traffico permettono di capire effetti di DOS in atto. Gli accorgimenti configurazionali di ntop sono in funzione della sezione di traffico che si vuole monitorare. Per grandi flussi con un numero di connessioni molto alte è preferibile una configurazione a flow dato che lo strumento in modalità sniffing analizza in real time il traffico caricando la macchina. Le modalità di inserimento di un dispositivo per l'analisi del traffico devono contemplare, come per IDS, la replica del traffico tramite una span port, mirror port, TAP o traffic replication.

In strumenti così variegati come ntop è necessario affidarsi all'effetto ottico che i grafici rappresentati danno, confrontare i picchi con gli stessi periodi temporali nelle vari giornate, aumenti molto veloci di tipologia di traffico aumento spedito di sessioni contraddistinguono una anomalia, poi l'operatore è in grado di interpretarla di modo da compiere delle azioni tramite strumenti di pre-

IRC report data, 30-second sample, then summarizations

[irc summarization](#) for last sample period (last 30 secs) (ASCII)

daily summarizations of irc stats.

[irc summarization for today \(run hourly\)](#) [yesterday](#) [today - 2 days](#) [today - 3 days](#) [today - 4 days](#) [today - 5 days](#)

RRDTOOL global network irc message count data

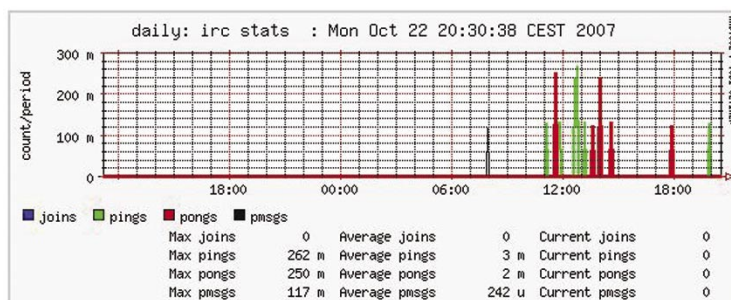


Figura 4. XXXXXXXXXXXXXXXXXXXX[Fig 4 ourmon1.jpg]

venzione o di filtering.

Ourmon

Ourmon è un analizzatore statistico di rete; il suo scopo è quello di isolare in tempo reale le informazioni rilevanti dal "rumore" di fondo della rete. La modalità di funzionamento per la quale è stato concepito è l'analisi del traffico (ottenuto usualmente tramite port mirroring) con ispezione dei pacchetti fino al livello 7 della pila ISO/OSI; non sono supportati nè lo storing dei dati "non filtrati", nè il monitoraggio di flussi nFlow.

I primi sono infatti considerati inutili secondo la filosofia di ourmon, secondo la quale l'informazione grezza è valida solo per il momento presente e solo al fine di costruire statistiche su poche variabili pregiate, che meritano quindi di essere conservate. L'utilizzo delle informazioni nFlow è invece inopportuno proprio sulla base delle modalità operative, che prevedono di elaborare statistiche partendo dai *dati reali*, e non da elementi già sintetizzati; del resto, sono presenti in letteratura molte istanze favorevoli all'approccio *packet-based* contro quello *flow-based*.

Le critiche principali che sono mosse a quest'ultimo sono:

Incapacità di cogliere fenomeni sporadici e spike.

Ritardo nell'identificazione delle anomalie, in quanto i dati nFlow sono elaborati (usualmente da router) con uno specifico periodo.

Introduzione di inesattezze nell'eventualità in cui, per alleggerire il carico di CPU dei router, i dati nFlow si basano sul campionamento dei pacchetti, limitando l'analisi a 1 pacchetto ogni n; questo ha anche l'indesiderato effetto collaterale di filtrare del tutto fenomeni troppo sporadici, che scompaiono dal rilevamento..

Dal punto di vista del software design, la struttura di Ourmon è modulare; come mostrato nella figura 4. una Probe Box, preposta alla raccolta di dati grezzi. (vedi figura5).

una Graphic Box, la quale, rispondendo ad una schedulazione cron, elabora automaticamente i dati estratti, aggiorna i relativi grafici e

personalizza i template HTML che verranno poi resi disponibili dall'interfaccia Web di reporting.

Tale interfaccia si presenta con un quadro riassuntivo delle variabili di rete prese in esame; queste ultime, ed i relativi grafici (generati dinamicamente da tool standard open source), sono specificate mediante un linguaggio di scripting; l'utente ha la possibilità di creare i propri filtri usando espressioni regolari scritte secondo la sintassi BPF (Berkeley Packet Filter) o PCRE (Perl Compatible Regular Expressions). Una volta evidenziato il traffico d'interesse mediante opportuni script, OurMon dà la possibilità di discriminare gli host che ne generano la maggior parte ("top talkers").

Una delle rilevazioni più utili per scovare eventuali botnet annidate nella propria rete è quella relativa al traffico IRC: dai canali IRC sono estratte le informazioni relative al numero di host attivi, di messaggi scambiati, di join e di host infetti da worm, e classificati sulla base della pericolosità stimata ("evil factor"). Ourmon infatti fornisce non solo i dati sull'attività del protocollo, ma correla tali informazioni con quelle relative agli scanner rilevati sulla rete: questo permette in automatico di marcare alcuni IRC host come compromessi ("evil hosts").

Operativamente, Ourmon supporta i sistemi Linux e FreeBSD; i requisiti hardware possono variare in funzione dell'ampiezza di banda da monitorare e del numero di filtri attivi. I log del sistema, non presenti nell'interfaccia web, possono dare utili indicazioni per verificare il corretto

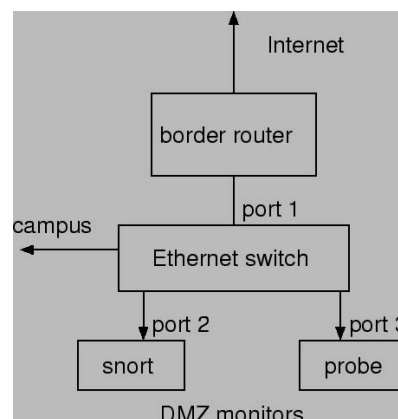


Figura 5. XXXXXXXXXXXXXXXXXXXX [Fig 5 ourmon2.png]

dimensionamento della macchina.

Esperienza operativa

Portiamo come esempio di utilità pratica di questi strumenti un caso reale: in una situazione di funzionamento normale, senza che gli strumenti IPS mostrassero risultati particolari, la quiete degli amministratori dei server viene turbata dalla comparsa di alcuni valori non nulli nel grafico Ourmon per IRC relativo alla server farm. Certi di non ospitare simili servizi, viene dato inizio ad una indagine: Ourmon marca tali host come "evil". La macchina in questione pare però funzionare senza anomalie; si notano però cambiamenti in alcuni file specifici e in alcuni utenti. La preoccupazione diviene reale, e con essa la possibilità di avere danneggiamenti. Viene rintracciata l'origine delle comunicazioni IRC: l'attaccante ha sostituito l'eseguibile SSH con una sua versione appositamente modificata; ripristinata la versione legittima, e visto il traffico IRC cessare, vi è più tempo per pensare a come

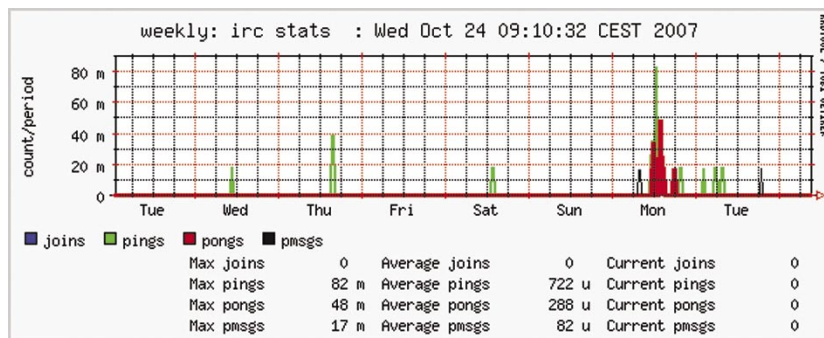


Figura 6. [Fig 6 ourmon3.png]



sia stato possibile bypassare i livelli di sicurezza. La risposta viene trovata nelle applicazioni in hosting sul server, facenti riferimento a librerie datate, talmente obsolete da essere state scartate dalle rules applicate dai layer IPS: quest'ultimo bloccava però taluni comandi che l'intruso cercava di lanciare, limitandogli i movimenti. Ripristinate le signature opportune, la quiete del grafico IRC diviene certezza che l'ospite indesiderato non ha ottenuto granchè dai suoi attacchi, se non una denuncia alla polizia postale.

Conclusione

Avere uno o più rilevatori di anomalie costituisce certamente un'arma in più per migliorare i livelli di servizio della propria infrastruttura di rete, aumentando la consapevolezza qualitativa sul traffico e garantendosi un ulteriore strato di difesa contro attacchi insidiosi o sconosciuti.

```
<RAMKAL posx=9;0l posy=b  
fit=W grow=H>>
```

In Rete

- <http://www.ntop.org/>
- <http://www.10t3k.org/security/doc/anomaly>
- <http://ourmon.sourceforge.net/>
- http://www.sagecertification.org/events/usenix05/tech/freenix/full_papers/binkley/binkley_html/index.html

Sull'autore

Snortattack.org, Portale orientato alla sicurezza informatica, è il risultato della fusione di conoscenze e collaborazione del team. Le tipologie di argomentazione trattate coprono 360 gradi tutte le tematiche relative alla sicurezza: attacco/difesa.

Grande punto di forza è l'uso di Snort come soluzione alle innumerevoli problematiche di intrusione. Per tenere gli utenti aggiornati sulle nuove problematiche sono a disposizione un forum e una mailinglist. Con Snortattack.org, si intende creare uno Snort User Group finalizzato alla collaborazione per l'Italia e tutto il resto del Mondo, per l'uso di Snort e la trattazione di problematiche di security.