La Gestione probabilistica del Rischio Cyber

A cura di: **Cesare San Martino**, IT Security Consultant, *EuroSel MoMa*

1. La definizione tradizionale di Rischio informatico ed i suoi limiti

Nella letteratura corrente relativa alla sicurezza informatica è uso definire il Rischio Informatico come:

$$R = V*T*I$$

Intendendosi che il Rischio IT è composto dai fattori:

V: Vulnerabilità dei sistemi
T: Threat, presenza di minacce in grado di far leva sulle vulnerabilità
I: Incident, costo dell'incidente qualora si realizzasse.

E' evidente l'estrazione "tecnica" di tale definizione. E' una definizione congegnata da sistemisti, preoccupati di dover "patchare" appropriatamente il proprio sistema, parte del loro lavoro quotidiano, per non esporre delle vulnerabilità delle quali potrebbero avvantaggiarsi malintenzionati minacciosi in attesa di fare l' "Exploit".

Fa parte della professionalità dei tecnici preoccuparsi di identificare e rimediare tutte le vulnerabilità ed essere al massimo grado al corrente delle minacce esistenti. Aggiungiamo che di regola il valore del bene da proteggere non rientra nei ragionamenti dei tecnici dell'IT Security.

Sostengo che tale definizione è **poco utile**; fuorviante se non direttamente falsa; di ostacolo alla comunicazione dello stato della sicurezza al business; inibente rispetto alla necessità di quantificare gli investimenti in IT Security.

Quali sono le vulnerabilità dei sistemi? Sappiamo tutti che ogni giorno ne escono di nuove. Sappiamo che esistono debolezze *zero day*, cioè non note in pratica a nessuno se non agli hacker. Inoltre quando si parla di vulnerabilità in questo modo si hanno in mente le patch che mancano, i certificati malformati, le porte Telnet aperte; tutti elementi tecnici dei sistemi da gestire. Ma ha senso avere in mente queste vulnerabilità

quando è ben noto che la vulnerabilità maggiore è quella data dagli essere umani ? Che aprono allegati che non dovrebbero aprire, che navigano dove non dovrebbero, che danno informazioni al telefono che non andrebbero date. In che modo quantifichiamo le vulnerabilità umane, che sono pur sempre le maggiori?

Quanto alle minacce, di nuovo, come catalogarle? Come quantificarle? Come facciamo a sapere se esiste una nuova minaccia in grado di attaccare una vulnerabilità forse nota e forse no?

Infine l'Incident, il terzo elemento dell'equazione. Si intende qui il "costo" dell'incidente. Esso è di regola valutabile dalle aziende, anche se può non sempre essere semplice, dipendendo dal valore del sistema ma anche da elementi immateriali quali la perdita di reputazione o altro.

In sintesi mi sembra che l'accoppiata Vulnerabilità / Minaccia sia una **metafora ben poco praticabile**. Essa immagina un modello **deterministico** in cui tutte le vulnerabilità e le minacce sono note ed i sistemi sono perfettamente gestiti in tempo reale. L'immagine che si crea è: "se rimedio tutte le mie vulnerabilità, se sono al corrente di tutte le minacce, i miei sistemi saranno perfettamente sicuri". Sappiamo bene che non è così. Non è così nelle organizzazioni complesse in cui il quadro vulnerabilità / minacce è semplicemente troppo vasto; e non è così nelle organizzazioni più piccole in cui la complessità sarebbe dominabile ma dove non esistono skill e risorse adeguate a gestirla.

2. La natura probabilistica del Rischio, in generale

Le aziende che vivono di Rischio, le assicurazioni, usano un modello ben più semplice e significativo di Rischio:

$$R = p*I$$

Questa è peraltro anche la definizione intuitiva che ciascuno di noi effettua in mille occasioni quotidiane:

- Assicuro la macchina contro l'incendio? So che la probabilità è bassa, dipende dal valore della macchina.
- Mi sveglio un'ora prima per prendere il treno prima o dormo di più prendendo il treno dopo se devo arrivare ad un incontro? Probabilmente ho esperienza della tratta e so mediamente quanto in ritardo è il treno, ma dipenderà anche dall'importanza dell'incontro.

E basta andare su Wikipedia per trovare: "Il Rischio è definito dal prodotto della frequenza di accadimento e della gravità delle conseguenze".

Tornando alle assicurazioni (certamente il tipo di impresa di gran lungo più maturo in tema di gestione del Rischio), sappiamo che la parte "difficile" dell'equazione è la probabilità di un accadimento. Ma questa aziende dispongono di potenti modelli matematici e di statistiche storiche di decenni se non di secoli in alcuni casi, che permette loro una valutazione sensata della probabilità. E' nota peraltro la **riluttanza** delle assicurazioni ad assumersi Rischi, anche molto ben remunerati, in assenza di una capacità di valutare la probabilità di un accadimento.

Ben diversa quindi la capacità di valutazione del rischio in confronto ad elementi in teoria determinati ma in pratica difficili da identificare quali vulnerabilità e minacce.

Avvicinandosi al nostro settore vediamo come oggi le aziende che assicurano il Rischio Cyber, ad oggi principalmente americane, utilizzano **modelli probabilistici** nell'assunzione del Rischio Cyber.

Sembra in definitiva intuitivo dire che la **natura del rischio è probabilistica**; che forzare il concetto di rischio in gabbie deterministiche non sia utile; ed infine che il business, quale che esso sia, ragiona il rischio in modo stocastico e quando gli specialisti di sicurezza informatica si relazionano con il business parlando di "vulnerabilità e minacce", il rischio che diventi un "dialogo tra sordi" è decisamente alto.

3. Come coniugare Rischio informatico e natura probabilistica del Rischio

Detto quanto sopra, come coniugare quindi la visione tradizionale e corretta di Rischio nel contesto IT, nel contesto Cyber?

Di passaggio notiamo come le organizzazioni più avanzate, soprattutto nel settore Finance, parlano di sicurezza IT in termini di Gestione del Rischio generale.

E' essenziale questo passaggio anche per razionalizzare gli investimenti in Cyber Security. Nel valutare se implementare un certo controllo va valutato il costo dell'implementazione rispetto al danno dell'incidente moltiplicato per la sua probabilità di accadimento. Solo in questo modo si possono prendere **decisioni economicamente razionali** sugli investimenti in sicurezza IT.

Come quindi valutare la probabilità degli incidenti? Ecco che qui che, cacciate dalla porta, possono dalla finestra utilmente rientrare misure come Vulnerabilità e Minacce. Dovremmo servirci di:

- Indicatori di vulnerabilità e minacce. Potenziali Vettori di Rischio.
- Modelli matematici / algoritmici che ci permettano di correlare i Vettori di Rischio per giungere a calcolare le probabilità di accadimento di certi eventi
- Se possibile, e nel corso del tempo, dati storici sulla probabilità di eventi negativi di sicurezza, organizzati magari per Industry / Nation etc.

Se quanto sopra sembra troppo teorico e poco praticabile guardiamo questo esempio, uno studio elaborato da una società commerciale evidenzia la forte correlazione tra presenza di macchine infette all'interno di una rete e probabilità di Data Loss: *Beware the Botnets: Botnets Correlated to a Higher Likelihood of a Significant Breach, BitSight Insights, April 2015* (http://www.advisenltd.com/wp-content/uploads/2015/04/beware-the-botnets-bitsight-paper-2015-04-23.pdf).

La parte essenziale:

	No Breach	Breach	All	Percent Breached
Botnet: A	1,538	26	1,564	1.7
Botnet: B or lower	4,536	172	4,709	3.7
All	6,074	199	6,273	3.2

mostra come le realtà con maggiore presenza di BOT hanno una probabilità più che doppia rispetto alle altre di essere soggette e fenomeni di Data Loss. La causa potrebbe farsi risalire alle Botnet stesse, ma non è detto. Non c'è una dimostrazione causale, deterministica dei motivi di Data Loss. C'è una correlazione probabilistica tra un certo Vettore di Rischio ed un accadimento negativo, fissato in una percentuale precisamente calcolata.

Potrà non essere semplice determinare quali elementi tecnici agiscono da vettori di rischio. Si potrebbero citare:

- Presenza di Botnet e Malware in generale
- Fenomeni di Spam
- Porte aperte su protocolli rischiosi
- Qualità dei certificati esposti
- Qualità dei protocolli Internet esposti
- Presenza di fenomeni legati a comportamenti utente rischiosi (downloads, torrent etc)

e altri ancora.

Ancora meno semplice sarà congegnare un algoritmo che “macini” opportunamente gli elementi sopra esposti, dia ad ognuno un diverso peso in relazione alla sua gravità ed elabori infine un indicatore del Rischio Cyber, l’agognata **p** di cui si parlava prima.

4. Conclusioni

- La visione deterministica del Rischio Cyber è poco utile, fuorviante; rende difficile il colloquio tra Business e IT; non mette in grado di valutare la convenienza degli investimenti in Sicurezza informatica.
- La visione probabilistica risponde naturalmente all’essenza del concetto di Rischio ed elimina gli inconvenienti sopra detti.
- Sarà utile elaborare indicatori di rischio ed algoritmi per trattarli che indichino la probabilità di accadimenti negativi. Esistono offerte commerciali che basandosi su Open Data permettono tale approccio, prima tra tutte BitSight Technologies.