ANALISI, VALUTAZIONE DEL RISCHIO E SICUREZZA INFORMATICA DI DATI E INFORMAZIONI DEI DISPOSITIVI MEDICI CONNESSI ALLE RETI IT-MEDICALI

C. Chierchia[1], E. Guerra[3], M. Balloccu[3], F. Deluca[1], L. Monasta[2], M. Bava[1]

[1]Ufficio Sistema Informativo – SC Ingegneria Clinica, Informatica e Approvvigionamenti – IRCCS materno-infantile "Burlo Garofolo" – Trieste
[2]SSD Epidemiologia e Statistica – IRCCS materno-infantile "Burlo Garofolo" – Trieste
[3]DIA – Università degli Studi di Trieste - Trieste

L'attuale evoluzione tecnologica ha permesso di collegare e far convergere due mondi fino ad almeno quindici anni fa sconnessi tra loro e poco comunicanti: quello biomedicale, in particolare dei dispositivi medici (DM), e quello IT relativo alla sicurezza informatica.
Se fino ad ora questa evoluzione ha portato all'interno degli ospedali un netto miglioramento nel processo e nel trattamento dei dati in termini di tempo (es. archiviare un referto) spazio (es. magazzini dove vengono conservate le cartelle cliniche) ed, almeno in parte, anche in termini economici, adesso non è possibile non considerare come fondamentale, viste le enormi quantità di dati clinici e sanitari che gestiscono i sistemi, il punto di vista della sicurezza.
La mancanza di idonei strumenti di protezione della rete dati ospedaliera può portare non solo al danneggiamento della macchina e di conseguenza dei pazienti, ma anche rendere la rete stessa facile preda di *hackers* malintenzionati che, ad esempio con *malware* come *cryptolocker* e affini, possono ricattare gli enti in cambio dello "sblocco del sistema" o peggio ancora recuperare illegalmente dati personali e sensibili vendendoli al miglior offerente.
Nella attuale società della conoscenza possiamo considerare la protezione dei dati e le informazioni che essi generano (*data and system security*) come il nuovo oro nero, cambiando e innovando il modo di produrre ricchezza. In Europa per esempio le aziende che commercializzano prodotti o servizi relativi ai dati sono circa 255.000, producendo quasi il 2% del PIL e con la prospettiva futura di arrivare nel 2020 al 4% del totale dell'economia dell'Unione Europea. Inoltre, secondo la direzione generale delle Reti di Comunicazione, dei Contenuti e delle Tecnologie (Dg Connect) della Commissione Europea, nel 2016 il valore del mercato dei dati in UE è stato di 59.53 miliardi di euro, in crescita rispetto all'anno precedente (54.3 miliardi di euro) e le previsioni per il futuro

sono ancor più rosee o inquietanti: si pensa infatti che nel giro dei prossimi tre anni il valore possa superare i 106 miliardi di euro.

Ora, sebbene sia difficile fare una stima del valore dei dati personali venduti illegalmente, alcuni parametri ci permettono di capire che la questione è sentita e per nulla scontata: dal 2013 al 2016 il nucleo speciale Privacy della Guardia di Finanza ha comminato sanzioni fino a un valore massimo di 400 mila euro e per un totale complessivo di circa 20 milioni[1]. Inoltre assistiamo oramai a un aumento esponenziale di attacchi sempre più complessi e articolati delle reti informatiche da parte di *cyber*-criminali che realizzano gran parte del loro fatturato con la sottrazione e la violazione di dati di aziende impreparate ad affrontare efficacemente la minaccia[2].

È quindi fondamentale e auspicabile che gli enti si tutelino con misure di sicurezza adeguate per proteggere capitale, tecnologia e conoscenza, in particolar modo i sistemi e i servizi che trattano dati sensibili, attraverso investimenti sulla messa in sicurezza della rete informatica e dei sistemi stessi (DM, apparecchiature, ecc…) che producono questi dati ed informazioni. Lo scopo è quello di ottenere un elevato grado di protezione da attacchi esterni e garantire così la continuità operativa della struttura. Ma non basta: è necessario infatti che questo grado di protezione sia periodicamente verificato attraverso un'analisi del rischio che produca parametri oggettivi per la valutazione di tutti i sistemi, servizi e le apparecchiature collegate alla rete IT-medicale[3].

Nel progetto che proponiamo, partendo da quanto acquisito in ambito legislativo (D.Lgs. 196/03, nuovo Regolamento Privacy 679/2016), e normativo (ISO 27001, ISO 80001 e ISO 30001) si attribuisce, ad ogni apparecchiatura o dispositivo medico collegato ad una rete ospedaliera, un indice, l'Indice di Valutazione del Rischio (IVR), che valuti il relativo livello di sicurezza nelle condizioni d'uso tipiche. L'IVR viene ottenuto attraverso l'implementazione di metodi statistici e una stima dei pesi oggettiva che renda il modello ripetibile e quindi convalidabile.

L'IVR è distribuito in un *range* di valori da 1 (basso rischio) a 10 (alto rischio) e suddiviso in macrocategorie che tengano conto sia delle tematiche tipiche dell'ingegneria clinica (la documentazione e la manutenzione delle apparecchiature, i rischi collegati al paziente) sia di aspetti ICT solitamente trascurati nell'analisi del rischio delle tecnologie biomediche. Attraverso l'assegnazione di una serie di regressori viene realizzata la formula per il calcolo dell'IVR relativo al rischio rilevabile sulla singola apparecchiatura nelle condizioni di esercizio.

In particolare per la parte relativa alla sicurezza informatica si è scelto di considerare come regressori la presenza/assenza di credenziali di accesso per accedere al sistema, se è presente ed è aggiornato l'antivirus, se è stato effettuato il backup dei dati, se è avvenuta perdita recente dei dati,

se è attivo un firewall, se il dispositivo è sotto gruppo di continuità e se risulta positivo ai test di vulnerabilità: tutti argomenti che riguardano la privacy, la *information security* e la *cybersecurity*.
Per ricavare i pesi di ciascuna categoria, nel nostro studio sono stati utilizzati due modelli allo scopo di confrontarne i risultati: la regressione lineare multipla e il modello logistico. Nel primo caso si studia la dipendenza di una variabile quantitativa Y dall'insieme dei regressori $X_1,\ldots,Xm$, $Y = f(X_1, \ldots, X_m) + \varepsilon = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m + \varepsilon$ attraverso un modello lineare[4], e una tripartizione del livello di rischio in alto, medio e basso; nel secondo caso invece il modello di regressione è applicato nei casi in cui la variabile dipendente Y può assumere esclusivamente valori dicotomici, in questo caso alto e medio/basso rischio $Y=logit(p) = ln\left(\frac{p}{1-p}\right)$.
I risultati finora ottenuti evidenziano che entrambi i modelli possono essere presi in considerazione ed essere valutati per la stima dei pesi per le singole categorie e quindi trovare l'IVR dell'apparecchiatura; il modello logistico però riesce a simulare meglio l'andamento desiderato con una sensibilità del 100% e una specificità dell'82%, identificando tutti i dispositivi ad alto rischio, mentre per il medio/basso rischio non individua quattro macchine, ponendole ad un livello di allerta superiore al necessario.

Avendo un modello ripetibile e convalidabile la prospettiva futura è di impiegarlo nelle strutture ospedaliere per fornire una valutazione del rischio realistica e affidabile, utilizzando una formula predittiva che permetta l'intervento tempestivo sul dispositivo e sulla rete dati, riducendo così i possibili rischi legati ad attacchi informatici o relativi allo stato delle apparecchiature. Non solo: l'espansione del suo utilizzo a livello territoriale permetterebbe di centralizzare l'archiviazione dei dati relativi alle strumentazioni biomediche delle strutture ospedaliere di tutta una regione.
Oltre che come inventario, l'aggregazione di una grande quantità di dati consentirebbe di applicare il modello descritto ai *big data* e ottenere così sia risultati dell'IVR sempre più attendibili, ma anche un sensibile miglioramento nell'attività decisionale. Attraverso lo studio di tali dati si potrebbero, infatti, analizzare eventuali variazioni dell'IVR e trovare i *trend* che permettano di prevenire i guasti, garantendo un ciclo di vita più lungo delle macchine e/o inviare un *alert* alle Aziende Sanitarie interessate prima che la criticità diventi troppo importante.
Inoltre, dagli stessi dati si può realizzare una valutazione statistica dello stato di salute dei dispositivi medici, sia considerandoli singolarmente, sia considerando complessivamente la struttura: un IVR medio molto discordante fra strutture analoghe appartenenti ad Aziende Sanitarie diverse potrebbe essere sintomo di una non adeguata manutenzione e controllo delle apparecchiature stesse.

L’utilizzo di reti neurali o altri sistemi di intelligenza artificiale o *machine learning* inoltre, permetterebbe di valorizzare ulteriormente la bontà del modello, con un supporto immediato e ancora più efficace alle decisioni in fase di analisi e valutazione del rischio.

[1] Lelio Simi: *“I dati sono il nuovo petrolio”* - http://www.pagina99.it/2017/03/16/industria-dei-dati-italia-nuovo-petrolio

[2] Laboratorio Nazionale CINI di Cybersecurity / Consorzio Interuniversitario Nazionale per l’Informatica 2016 *“Italian Cybersecurity Report”* pag.3

[3] CLUSIT – Associazione italiana per la sicurezza informatica; “Rapporto CLUSIT 2017 sulla sicurezza ICT in Italia”, 2017

[4] Montanari A.: La regressione lineare multipla – pag 1 - (http://www2.stat.unibo.it/montanari/Didattica/dispensa2.pdf)