Programma Operativo Nazionale 2014-2020
Dottorati di ricerca su tematiche dell'innovazione e green
D.M. n. 1061 del 10 agosto 2021

Università degli Studi di Napoli Federico II
Dottorato di Ricerca in Information Technology and Electrical Engineering
XXXVII CICLO
TEMATICHE GREEN (AZIONE IV.5)

BORSA N. 3 - Dottoranda SIMONA DE VIVO

Titolo del progetto:

"Augmented AI for Sustainable Cyber Security in Railway Environment"

SETTORE SNSI: Industria intelligente e sostenibile, energia e ambiente
SETTORE PNR: Sicurezza per i Sistemi Sociali; area di intervento: 5.3.3 Cybersecurity - Articolazione 1. Intelligence and incident response

a. Motivazioni

La sicurezza informatica, o ***cyber security,*** *è* intesa come l'insieme di metodologie e tecnologie volte a proteggere i sistemi, le reti e i programmi dagli attacchi digitali: oggigiorno rappresenta uno dei temi più di interesse sia nel mondo accademico sia nell'industria.
Secondo un report di Trend Micro [1], che ha raccolto dati nei primi 6 mesi del 2021, l'Italia è la prima in Europa in quanto attacchi malware, nel mondo invece è al quarto posto. Ovviamente la pandemia di Covid-19 ha peggiorato la situazione, dato che la migrazione massiva verso mezzi digitali in tutti i campi (come banking, pubblica amministrazione, sanità) ha esteso a dismisura la superficie d'attacco. Se questi attacchi prendessero come target un'infrastruttura critica, i possibili danni sarebbero decisamente ingenti.
Come specificato dal *Programma Nazionale per la Ricerca* (PNR), l'avanzamento della conoscenza scientifica nel campo delle metodologie e delle tecnologie orientate a prevenzione, identificazione, gestione, contenimento, analisi di attacchi cyber è tra le principali priorità, anche al fine di consolidare strutturalmente le capacità del Paese di difesa nel dominio cyber.
La cyber security oggi ha una ulteriore e nuova sfida: oltre a fronteggiare i sempre più complessi attacchi informatici, si deve adeguare anche al rispetto delle risorse dell'ambiente digitale, aprendo una nuova frontiera di ***cyber security sostenibile***.
Ogni giorno, le fughe di informazioni creano "*dump* di dati", ovvero fughe di informazioni su login e password nel dark web, che possono contaminare il nostro ambiente di dati (spesso noti in letteratura come *discariche di dati* o *dati spazzatura*). La soluzione per proteggere i dati risiede nella creazione di un ambiente di sicurezza informatica sostenibile, parallelamente alla linea di sostenibilità ambientale, per proteggere continuamente le informazioni.

La sostenibilità ambientale si concentra sullo sviluppo economico, sociale e ambientale. Evitando l'esaurimento delle risorse, possiamo continuare a nutrirci ed espandere le attività in modo responsabile. La sicurezza informatica sostenibile si declina nell'ambiente digitale. Il crescente interesse per il tema della cyber security, soprattutto nel presente periodo di forte trasformazione digitale, e l'amplificazione delle raccomandazioni sull'adozione di misure di prevenzione e mitigazione degli attacchi informatici (si veda ad esempio la legge n.133 dl 2019, nota come perimetro di sicurezza nazionale cibernetica) che portano ad interrogarsi su quale sia l'impatto sulla sostenibilità ambientale di siffatte raccomandazioni.
Un articolo dell'Università dell'Illinois [2] spiega l'importante ruolo svolto dal ***monitoraggio continuo*** della sicurezza informatica per la creazione di una cyber security sostenibile. L'articolo sostiene che, allo stesso modo in cui i fiumi diventano inutilizzabili a causa della pesca eccessiva, i messaggi di spam di un utente malintenzionato inquinano un ambiente di dati tramite un attacco DDoS (*distributed-denial-of-service*) o il *phishing* esaurendo la larghezza di banda limitata.
In altre parole, il mantenimento di un ambiente sostenibile - fisico e digitale - richiede una soluzione di monitoraggio proattivo dell'ambiente per garantire che le risorse rimangano accessibili. Diventa quindi fondamentale orientare la ricerca nazionale verso lo sviluppo di metodologie e tecnologie di monitoraggio per identificazione e contrasto delle attività di cybercrime. Un capitolo specifico è quello dell'analisi, della classificazione e del rilevamento di attacchi. L'analisi di attacchi e le strategie di ***attack detection*** rientrano sicuramente nel dominio più vasto dell'*incident response*, intesa come capacità delle organizzazioni di identificare, classificare, dare priorità, analizzare prontamente attacchi informatici, e attuare strategie di contenimento e *remediation* per un'analisi post mortem dell'incidente e l'estrazione di evidenze e prove digitali.
Il contesto del progetto sono le infrastrutture ferroviarie, sistemi di elevata importanza nazionale tanto da essere considerate vitali in quanto costituiscono le infrastrutture critiche nazionali (ICN). Le ICN sono infatti tutte le risorse, processi e sistemi davvero necessari al funzionamento di una nazione, come sicurezza, economia, telecomunicazioni, sanità, e, appunto, trasporti. È proprio per la loro importanza che sono potenziali target di attacchi informatici.
La gestione dell'infrastruttura ferroviaria svolge un ruolo importante per rispondere alla pressione commerciale richiesta dal trasporto nazionale. Infatti, sia le aspettative della società per salvaguardare la sicurezza e la salute delle persone e sia le esigenze del mercato per l'efficacia dei costi e il livello di servizio stanno diventando sempre più importanti nel nuovo contesto competitivo derivante dalla liberalizzazione dei trasporti ferroviari in Europa.
Molte aziende ferroviarie devono soddisfare le regole previste dalle normative di sicurezza, che, in diversi Paesi, definiscono le modalità di manutenzione e anche le frequenze per la manutenzione preventiva con l'obiettivo primario di fornire un elevato livello di sicurezza. Inoltre, la ferrovia oggi è in concorrenza con altre forme di trasporto e le società ferroviarie vengono suddivise per fornire servizi di trasporto da un lato e servizi infrastrutturali

dall'altro per diventare più efficaci. Inoltre, i clienti desiderano sempre più un servizio di migliore qualità a costi inferiori, costringendo le aziende ferroviarie a ottimizzare i costi in ogni fase del processo, compresa la manutenzione.
Questa situazione crea un contesto particolarmente sfidante per fornire sicurezza alle infrastrutture ferroviarie, data l'enorme varietà di elementi tecnologici installati in un'infrastruttura ferroviaria, rendendo così molto complessa la governance delle risorse coinvolte nelle operazioni di manutenzione. Tali infrastrutture, infatti, sono oggigiorno delle reti complesse che contengono centinaia di nodi comunicanti con diverse tecnologie (Wi-Fi, wired, ecc.) per effettuare diversi compiti, quali il controllo degli attuatori, per parlare con la stazione esterna, ecc.
Nel contesto sopra descritto, le imprese ferroviarie stanno compiendo sforzi significativi per aumentare la sicurezza delle loro infrastrutture, con l'obiettivo di ridurre i costi operativi mantenendo elevati standard di sicurezza. Inoltre, hanno iniziato ad esternalizzare alcuni servizi di manutenzione per testare, in confronto, la correttezza delle loro procedure di manutenzione. Ad esempio, Rete Ferroviaria Italiana (RFI), la società per azioni italiana per il trasporto ferroviario, ha avviato ricerche per migliorare le proprie procedure di testing, cercando di correlare meglio le politiche di manutenzione applicate con l'affidabilità e la sicurezza raggiunta dell'infrastruttura.

b. Obiettivi

Seguendo le linee guida dettate dal PNR, e in particolare l'area di intervento *5.3.3 Cybersecurity - Articolazione 1. Intelligence and incident response,* il progetto si pone **l'obiettivo di sviluppare una soluzione per proteggere, identificare e rispondere ad attacchi di cyber security attraverso una soluzione di monitoraggio delle infrastrutture che sia automatica, a basso consumo energetico e a basso consumo di risorse di elaborazione**.
Le attività di ricerca del progetto sono da ritenersi in linea con le aree tematiche individuate dalla *Strategia Nazionale di Specializzazione Intelligente* (SNSI). In particolar modo, le attività di ricerca sono trasversali all'area tematica I*ndustria intelligente e sostenibile, energia e ambiente*, e alla traiettoria di sviluppo *Processi produttivi innovativi ad alta efficienza e per la sostenibilità industriale*.
Per la creazione della soluzione proposta, il progetto mira a sfruttare tecniche di *elaborazione del linguaggio naturale* (dall'inglese, *Natural Language Processing* - NLP). L'elaborazione del linguaggio naturale è un sottocampo della linguistica, dell'informatica e dell'*intelligenza artificiale* (IA) che si occupa delle interazioni tra computer e linguaggio umano, in particolare come programmare i computer per elaborare e analizzare grandi quantità di dati. L'obiettivo è quello di fornire un *engine* in grado di capire i contenuti dei documenti, comprese le sfumature contestuali del linguaggio al loro interno. La tecnologia può quindi estrarre con precisione le informazioni e gli approfondimenti contenuti nei documenti, nonché classificare e organizzare i documenti stessi. Le tecniche di NLP sono state usate nei più svariati contesti, come il monitoraggio dei social media per identificare

crimini ambientali (ad esempio, combustione illecita, rifiuti radioattivi, discarica abusiva, ecc.), analisi di log testuali per identificare anomalie, identificazione di potenziali abusi di diritti umani nei flussi di dati continui dei social media, ecc.

Il progetto mira a sfruttare l'intelligenza artificiale, e in particolare le oramai avanzate tecnologie di NLP, per effettuare monitoraggio delle infrastrutture, al fine di individuare eventuali anomalie dovute ad attacchi informatici, implementando di fatto una soluzione intelligente di analisi e detection degli attacchi che possa garantire le risorse sempre disponibili, evitando fughe di dati dovute ad attacchi e quindi inserendosi a tutti gli effetti nel contesto della cyber security sostenibile.

Il monitoraggio delle infrastrutture permetterà anche di raccogliere le informazioni sugli attacchi (***cyber threat intelligence***), che sono utilizzate per capire quali solo le minacce informatiche che prendono, avevano preso o prenderanno di mira le infrastrutture, e sono informazioni essenziali per identificare e prevenire tali attacchi.

I modelli di *deep learning* sono oramai la prima scelta quando si parla di NLP e più in generale di IA, dato che spingono sempre più in alto i risultati dello stato dell'arte in diversi campi e contesti. Tuttavia, nonostante i risultati promettenti, bisogna constatare che questi modelli richiedono calcoli massicci per ottenere risultati migliori, e infatti le risorse computazionali richieste sono aumentate a un ritmo esponenziale. Basti pensare che la quantità di calcoli utilizzati per addestrare i modelli di deep learning è aumentata di 300.000 volte in 6 anni [3]. Questi calcoli non solo causano costosi costi finanziari, ma contribuiscono anche a un'emissione di carbonio sorprendentemente elevata. Il primo danneggia l'inclusività dell'IA nel mondo accademico e dell'industria, il secondo danneggia il nostro ambiente.

Si parla formalmente di ***Green AI***, che fa appello ai ricercatori per ottenere nuovi risultati senza aumentare il costo computazionale, ma riducendolo idealmente [3]. A differenza di *Red AI* che spinge a tutti i costi i risultati dello stato dell'arte, il Green AI incoraggia i ricercatori a ottenere risultati comparabili o - se possibile - migliori utilizzando il minor numero di risorse computazionali possibile. Quando si parla di metriche computazionali, la comunità scientifica fa riferimento a tempo di esecuzioni, emissioni di carbonio, dimensioni dei modelli di *deep learning*, operazioni di *floating point*, ecc.

Il design della soluzione di monitoraggio che il progetto mira a sviluppare si pone nel nuovo contesto della Green AI, in quanto si propone di rispettare le seguenti caratteristiche:

- **Architettura Compatta**: Costruire modelli compatti, cioè reti neurali con dimensioni ridotte (numero di layers, numero di neuroni, ecc.) in grado di ottimizzare il rapporto performance/consumi.
- **Addestramento Efficiente del Modello**: Addestrare i modelli di deep learning in modo efficiente è fondamentale per ridurre i consumi energetici. Esistono varie strategie per l'addestramento efficiente della rete neurale, come l'inizializzazione, la normalizzazione, il training progressivo e AutoML efficiente.
- **Inferenza Energeticamente Efficiente**: Anche la predizione dell'output da parte del modello può essere ottimizzata in termini energetici. Una soluzione è ridurre il

numero di parametri ridondanti che influiscono poco o nulla sul miglioramento dei risultati (*pruning*).

- **Utilizzo Efficiente dei Dati:** È importante utilizzare in modo efficiente i dati usati per addestrare i modelli di deep learning. L'utilizzo di modelli pre-addestrati o di tecniche di *transfer learning* sono soluzioni più che valide in tale contesto.

L'IA ha un ruolo fondamentale nella soluzione proposta. Essa è già famosa come un'ottima soluzione ai problemi di sicurezza, ma non è l'unico vantaggio che apporta. Tramite l'IA, infatti, si può apportare un livello di automazione superiore, rendendo l'infrastruttura smart e riducendo così i tempi e i passaggi necessari per la sicurezza della struttura. Inoltre, rendendola anche green, si diminuisce l'utilizzo di risorse, come quelle energetiche, rendendo le aziende sempre più ecologiche e sicure, che risulta essere fondamentale in un contesto embedded come quello ferroviario.
Per testare la robustezza della soluzione di monitoraggio nella detection e classificazione degli attacchi stessi, si valuterà l'accuratezza in termini delle metriche comunemente usate nel contesto di *anomaly detection* (e.g., *true/false positive/negatives rates*). Per valutare i miglioramenti apportati dal punto di vista energetico e computazionale, si valuteranno le riduzioni di emissioni di carbonio, di consumo energetico, di tempi di esecuzione, di complessità dei modelli realizzati, ecc. Il gruppo di ricerca proponente non è nuovo a questo tipo di validazioni, come si evince da lavori scientifici precedenti [4].

c. Cronoprogramma

Il progetto si articolerà su un periodo temporale di 36 mesi, durante il quale il dottorando studierà lo stato dell'arte delle tecnologie attuali, e progetta e sviluppa le soluzioni innovative. La seguente tabella riassume il programma del progetto, suddiviso per anni.

Anno	Attività e Finalità
1	Fase di progettazione e configurazione per collezionare dati dell'infrastruttura
	Raccolta dati, studio minacce e creazione cyber threat report in ambito ferroviario
	Studio e individuazione di indicatori di consumo energetico ed emissioni di carbonio applicati nel contesto dell'intelligenza artificiale
2	Studio tecniche di *natural language processing* presso l'istituzione estera e applicazione a casi di studio pratici, effettuando la sperimentazione sul *testbed* disponibile presso l'università ospitante
	Applicazione delle tecniche di *natural language processing* per creare una soluzione di monitoring e *attack detection* automatica
	Riduzione del consumo energetico e delle risorse di elaborazione della soluzione della soluzione di monitoraggio e *attack detection*

3	Fase di analisi e sintesi dei risultati
	Misurazione e validazione del risparmio energetico e computazionale
	Periodo di formazione svolto presso l'azienda

d. Risultati attesi

La timeline del progetto include quattro **milestones**. Il raggiungimento di singoli traguardi consentirà di acquisire conoscenze sull'andamento delle attività del progetto. La tabella di seguito descrive i milestones, i tempi e i metodi di verifica.

Milestone	Nome	Anno	Metodo di Verifica
1	*Creazione cyber threat report*	1	Un documento e un repository che elencano le informazioni sugli attacchi alle infrastrutture ferroviarie
2	*Creazione soluzione di monitoraggio basata su NLP*	2	Rilascio su una repository pubblica di una soluzione di monitoraggio basata su NLP in grado di identificare anomalie dovute ad attacchi
3	*Valutazione accuratezza della soluzione di monitoraggio*	3	Un documento che elenca i risultati sperimentali condotti sulla soluzione di monitoraggio in termine di accuratezza nella detection degli attacchi
4	*Validazione risparmio energetico e computazionale della soluzione di monitoraggio*	3	Un documento che valida il risparmio energetico e computazionale della soluzione di monitoraggio, in tempo di tempi di esecuzione, di complessità della soluzione, di consumo di energia e emissioni di carbonio

e. Impresa

L'attività di ricerca sarà svolta in collaborazione con l'azienda Hitachi Rail STS, un player globale nel settore ferroviario: un gruppo integrato capace di offrire veicoli per il trasporto ferroviario, sistemi di segnalamento e tecnologia digitale, attività di Service & Maintenance nonché soluzioni chiavi in mano in tutto il mondo.

L'azienda è attiva in circa 30 paesi, con circa 4.000 dipendenti ed ha sede a Genova. Opera nella progettazione, realizzazione e gestione di sistemi e servizi di segnalamento e supervisione del traffico ferroviario e metropolitano, anche come lead contractor.

Negli ultimi anni, il settore della sicurezza informatica sta vivendo una evoluzione tecnologica legata alla crescente richiesta di sicurezza dal mondo dell'industria. In questo

contesto le attività di ricerca e risultati conseguiti, renderanno lo studente una figura professionale di notevole interesse applicativo per le aziende.
Gli argomenti affrontati dal progetto sono tutti associati a una domanda di mercato in costante aumento. Per questo motivo, il profilo del potenziale studente sarà decisamente interessante sia per le aziende che erogano servizi in diversi settori. Nel suo iter formativo lo studente, inoltre, si arricchirà nel dialogo con la comunità scientifica internazionale sulle tematiche più attuali della Ricerca e Sviluppo (R&D), maturando familiarità verso le nuove tecnologie e la loro ricaduta sul sistema produttivo, delineandosi come una figura strategica capace di coniugare sinergicamente Accademia e Impresa. Lo studente sarà guidato lungo un percorso di alta formazione tecnico-scientifica che gli garantiranno strumenti metodologici e competenze tali da delineare una professionalità altamente qualificata nell'ambito della ricerca e dell'innovazione tecnologica.
La capacità di gestire l'innovazione e l'originalità creativa nell'elaborazione e nella realizzazione di progetti e servizi acquisite dallo studente saranno aspetti strategici che risulteranno particolarmente utili all'interno delle PMI che intendano investire in R&D. Lo studente al conseguimento del titolo si proporrà nel mondo del lavoro come un "problem solver" con alta qualificazione scientifica in grado di ricoprire profili professionali di elevato livello grazie alla capacità di affrontare in modo autonomo i problemi ed alla predisposizione a lavorare in gruppo, spesso in contesti di carattere internazionale.

f. Istituzione ospitante all'estero

L'attività di ricerca sarà svolta in collaborazione con l'Università del North Carolina a Charlotte (UNCC), Carolina del Nord, Stati Uniti d'America, dove è presente un gruppo di ricerca che lavora con le tecniche di elaborazione di linguaggio naturale, supervisionato dal prof. Bojan Cukic. Il dottorando lavorerà a stretto contatto con tale gruppo di ricerca per un periodo di 8 mesi, al fine di poter imparare e affinare l'applicazione di tecniche di NLP.
Il periodo all'estero consentirà allo studente di raggiungere un livello avanzato di esperienza nel campo dell'applicazione di tecniche di NLP. Lo studente avrà l'opportunità di partecipare a seminari, conferenze e incontri organizzati regolarmente dall'università ospitante, incrementando le proprie capacità di esposizione dei risultati scientifici raggiunti con i propri pari e con il personale docente. Inoltre, lo studente avrà l'opportunità di apprendere l'uso di approcci basati su NLP sviluppati da UNCC e applicarli a casi di studio pratici, effettuando la sperimentazione sul testbed disponibile presso l'università ospitante. Tale sperimentazione pratica permetterà dottorando di sviluppare le sue capacità nel trasferimento tecnologico delle soluzioni di ricerca su sistemi reali di elevata complessità.

g. Prodotti misurabili della ricerca, comunicazione e disseminazione

Come attestazione delle attività di ricerca, il progetto si pone l'obiettivo di pubblicare dei lavori scientifici a conferenze e riviste. In particolare, l'obiettivo è quello di presentare articoli scientifici che riportano idee e risultati a conferenze e riviste leader in una serie di

aree di affidabilità, sicurezza, networking, ingegneria del software e ricerca di sistema in generale. Le sedi di pubblicazione target per la nostra ricerca includeranno, ma non sono limitate a quanto segue: IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Knowledge Discovery and Engineering, IEEE/IFIP International Conference on Dependable Systems and Networks, IEEE International Conference on Software Engineering, IEEE International Symposium on Software Reliability Engineering. Inoltre, i prodotti sviluppati (codice, dati) saranno rilasciati online su repository pubbliche (e.g., GitHub). Infine, il progetto mira anche ad organizzare workshop e sfide con utenti reali. L'idea principale è che gli utenti utilizzino l'approccio per convalidare l'effettiva utilità delle soluzioni create, sia per la scrittura di attacchi software e sia per il monitoraggio attivo contro gli attacchi. I workshop saranno anche l'occasione per presentare il gruppo di ricerca e favorire lo scambio internazionale.

Riferimenti

[1] Trend Micro, 2020 Report on Threats Affecting ICS Endpoints, url: https://documents.trendmicro.com/assets/white_papers/wp-2020-report-on-threats-affecting-critical-industrial-endpoints.pdf?_ga=2.227048341.989013401.1637695378-314880634.1637695378

[2] Shackelford, Scott J., Timothy L. Fort, and Danuvasin Charoen. "Sustainable cybersecurity: Applying lessons from the green movement to managing Cyber Attacks." U. Ill. L. Rev. (2016): 1995.

[3] Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green ai. Communications of the ACM, 63(12), 54-63.

[4] Cinque, M., Cotroneo, D., Frattini, F., & Russo, S. (2015). To cloudify or not to cloudify: the question for a scientific data center. IEEE Transactions on Cloud Computing, 4(1), 90-103.