**Programma di dettaglio della tesi**

**(da presentare al completamento del 20% delle attività di tesi)**

(*compilato dallo studente entro due mesi dalla comunicazione dell’assegnazione della tesi da parte della Commissione Tesi a seguito dell’incontro con il docente relatore e del relativo avvio delle attività*)

Da inviare utilizzando il seguente form:

<https://forms.office.com/Pages/ResponsePage.aspx?id=22cHw9o91E2KTQl9IsuZ03nJw8jHdCpNpV5x3k_J3dRUOFBDOFRXVDNOWEVZNUIxWVNLWFI0MjJQVS4u>

**Nome Cognome studente**

Luigi Russo

**Docente relatore**

Giuseppe D’Aniello, Sabrina Senatore

**Insegnamento di riferimento**

Situation Awareness

**Titolo argomento di tesi**

Costruzione e specializzazione di un Knowledge Graph per la Cyber Situation Awareness

**Descrizione dell’attività**

1. ***Descrizione del problema che si vuole affrontare (circa 5 righe)***

Nella cybersecurity, l’analisi e la risposta a vulnerabilità e minacce sono essenziali per strategie di difesa efficaci, ma i metodi tradizionali spesso non reggono l’aumento dei dati e della complessità. La Situation Awareness consente una visione globale degli eventi in corso, cruciale per comprendere il profilo di sicurezza di un’organizzazione. Un approccio basato su knowledge graph per modellare mediante relazioni le minacce e vulnerabilità di sistemi, attraverso domande in linguaggio naturale, può aiutare gli utenti a identificare situazioni critiche, risolverle e rispondere efficacemente agli incidenti di sicurezza.

1. ***Analisi dello stato dell’arte, scientifico e/o tecnologico (circa mezza pagina)***

* *Se la tesi ha carattere prevalentemente progettuale – applicativo, identificare lo stato dell’arte tecnologico.*
* *Se la tesi ha invece carattere prevalentemente metodologico, fare riferimento a lavori scientifici o metodologie a cui la tesi è collegata.*
* *I riferimenti devono essere riportati nella forma [1] [2] ecc e devono essere inclusi nella sezione “Bibliografia”*

I grafi di conoscenza per la cybersecurity offrono approcci olistici per elaborare enormi volumi di dati complessi sulla sicurezza informatica provenienti da fonti diverse. Questi strumenti possono supportare gli analisti di sicurezza nell’ottenere informazioni dettagliate sulle minacce informatiche, migliorare la consapevolezza della situazione (cyber-situation awareness), ottenere nuove conoscenze, e comprendere le correlazioni tra i dati aggregati. Inoltre, permettono di visualizzare reti, flussi di dati e percorsi di attacco, agevolando la comprensione e la prevenzione delle minacce. Attraverso l’aggregazione e l’integrazione di dati eterogenei e la loro rappresentazione strutturata è possibile aumentare il livello di situation awareness consentendo il monitoraggio delle minacce in tempo reale e l’anticipazione di possibili scenari di rischio. In [1] si esaminano i modelli di dati basati su grafi più rilevanti utilizzati in questo ambito. Ci sono molte caratteristiche relative alla sicurezza informatica e ai processi di rete che devono essere memorizzate, e la semantica delle conoscenze catturate varia notevolmente a seconda del modello di dati del grafo utilizzato.

Nel contesto della cyber situation awareness (cyberSA), l’implementazione di sistemi di Retrieval-Augmented Generation (RAG) rappresenta un significativo avanzamento nelle capacità di analisi delle minacce informatiche. Rahman et al. [2] descrivono due approcci innovativi, CyRAG e GraphCyRAG, che integrano Large Language Models (LLM) con dati strutturati e grafi di conoscenza per fornire risposte accurate e ricche di contesto. CyRAG si focalizza sull’elaborazione di dati strutturati relativi alle entità CVE (Common Vulnerabilities and Exposures) [4] e CWE (Common Weakness Enumeration) [5], migliorando l’accuratezza delle informazioni contestuali fornite agli analisti. Al contrario, GraphCyRAG sfrutta un framework basato su grafi di Neo4j per analizzare in profondità le relazioni tra CVE, CWE, CAPEC (Common Attack Pattern Enumeration and Classification) [6] e ATT&CK (Adversarial Tactics, Techniques and Common Knowledge) [7], fornendo una visione integrata e interconnessa delle vulnerabilità e degli schemi di attacco.

Questi sistemi si distinguono per la capacità di integrare conoscenze dinamiche, basate su dati aggiornati in tempo reale con strumenti avanzati per il recupero delle informazioni. In particolare, GraphCyRAG [2] offre un’analisi dettagliata delle relazioni tra vulnerabilità, tecniche di attacco e possibili strategie di mitigazione, migliorando la comprensione dei potenziali vettori di attacco e delle loro implicazioni operative. I risultati preliminari dello studio evidenziano che l’integrazione dei grafi di conoscenza con sistemi RAG migliora significativamente sia l’accuratezza sia la profondità dell’analisi delle minacce, potenziando la capacità degli analisti di individuare minacce e prendere decisioni informate.

1. ***Finalità della tesi, contributo dello studente e descrizione dell’attività progettuale di tesi (circa una pagina)***

* *Descrivere dettagliatamente il progetto che si intende svolgere nell’ambito della tesi.* 
  + *Descrivere gli obiettivi, le metodologie/tecnologie adottate (che cosa lo studente farà durante la sua attività di tesi e come lo farà)*
  + *Dettagliare (se presente) il carattere di innovatività introdotto nel progetto rispetto allo stato dell’arte*
  + *Descrivere le fasi in cui si articolerà lo sviluppo del progetto di tesi*

L’obiettivo principale di questa attività di tesi è la modellazione e progettazione di una pipeline di Retrieval-Augmented Generation basata su Knowledge Graphs per la Cyber Situation Awareness.

In un contesto come quello della cybersecurity, in cui ci sono enormi quantità di informazioni complesse e veloci, un sistema del genere consentirebbe all’utente (es. analista della sicurezza) di raggiungere e mantenere un livello di Situation Awareness (SA) che gli permetta di identificare, comprendere e anticipare le minacce in evoluzione.

A tale scopo saranno utilizzati grafi presenti in letteratura o creati ad hoc, realizzati mediante modelli di knowledge graph embedding per supportare la codifica e il recupero di informazioni dalla base di conoscenza e un LLM esistente in letteratura per la rappresentazione della query in linguaggio naturale e generazione di risposte.

L’attività di tesi si articola nelle seguenti fasi:

* La prima fase riguarda un’analisi preliminare del dominio di interesse e l’analisi del relativo stato dell’arte. Verranno esplorate e comprese le metodologie e tecnologie esistenti nel campo della RAG, con particolare attenzione all’utilizzo di dataset relativi alla CyberSA e all’utilizzo di eventuali modelli di LLM impiegati per questi dataset ma comunque già disponibili in letteratura. Verranno individuate le best practices e le soluzioni più accreditate nell’applicazione di tali tecnologie.
* La seconda fase consiste nello studio dei dati e la raccolta di dataset rappresentativi del dominio di interesse. Inoltre, sulla base del dataset scelto e del contesto verrà realizzata una Goal Driven Task Analysis (GDTA) in modo tale da identificare gli obiettivi e le decisioni che possono consentire di incrementare il livello comprensione della situazione.
* La terza fase consiste nell’impiego del RAG basato su Knowledge Graph Embedding (KGE) model. A partire dal modello addestrato verranno estratte le rappresentazioni vettoriali delle entità e delle relazioni presenti nel grafo di conoscenza.
* Nella quarta fase la query effettuata dall’utente verrà codificata in un vettore utilizzando un Large Language Model già presente in letteratura.
* Verranno poi recuperati i documenti più rilevanti sulla base della query (fase di Retrieval). Questi documenti verranno poi utilizzati come contesto per la fase di Generation producendo una risposta contestualizzata e informata.
* Testing: Ci sarà una fase di test per valutare le prestazioni del sistema in condizioni realistiche, apportando miglioramenti e ottimizzazioni basate sui risultati dei test. In particolare, verranno valutate le performance sia in fase di Retrieval che in fase di Generation.

1. ***Descrizione del protocollo sperimentale o del setup sperimentale (circa mezza pagina)***

* *Sperimentazione che si intende effettuare (per tesi di tipo prevalentemente metodologico) o descrizione dettagliata del setup sperimentale (per tesi di tipo prevalentemente applicativo)*
* *Indici prestazionali che saranno utilizzati per misurare la bontà del lavoro svolto e riferimento ai valori di tali indici (se già disponibili) di sistemi esistenti o di metodi allo stato dell’arte*

Per valutare le prestazioni del sistema proposto, sarà definito un ambiente sperimentale che simula scenari realistici legati alla Cyber Situation Awareness. In particolare, verranno utilizzati dataset rappresentativi che saranno strutturati per creare un grafo di conoscenza che servirà come base per tutte le fasi successive.

Le query del sistema saranno progettate per rispecchiare esigenze operative degli analisti della sicurezza, includendo domande relative alla comprensione delle minacce emergenti, alla correlazione degli eventi e all’identificazione di misure di risposta adeguate.

Questo setting permetterà di verificare il funzionamento del sistema di RAG in condizioni che riflettono l’uso reale, analizzando le sue capacità di elaborare informazioni, effettuare retrieval mirati e generare risposte contestualizzate.

Il setup sperimentale comprenderà:

* Una fase di elaborazione e pre-processing dei dati iniziali per estrarre entità e relazioni, costruendo un grafo di conoscenza utilizzando approcci semantici.
* Una fase di modellazione di un KGE model sulla base del grafo di conoscenza costruito nella fase precedente.
* Implementazione di un sistema di Retrieval basato sulla similarità tra i vettori delle query e quelli presenti nel grafo, per identificare le informazioni più rilevanti.
* Utilizzo di un set di domande per analizzare le risposte del sistema con e senza contesto.
* Analisi qualitativa che consentirà di individuare la presenza di errori o informazioni mancanti.
* Per avvalorare le performance del sistema potranno essere utilizzate alcune tra le metriche di valutazione selezionate in base allo stato dell’arte, quali Accuracy, Precision, Hits@N e MRR.

1. ***Dettaglio dei dataset che saranno utilizzati per la sperimentazione (con riferimento al tipo di dataset, se reale o sintentico, al numero campioni, ecc) (circa 5 righe)***

*In tale sezione devono essere chiarite le motivazioni alla base della scelta di quel dataset evidenziando il fatto che il dataset sia rappresentativo di un’istanza reale del problema affrontato. Laddove non è possibile utilizzare un’istanza reale del problema, è necessario chiarire e motivare accuratamente le ragioni alla base di tale circostanza.*

Per la sperimentazione verrà utilizzato il dataset D3FEND [3], un knowledge graph creato per modellare e rappresentare contromisure di cybersecurity. Si tratta di un dataset reale che integra informazioni provenienti da brevetti, documentazione tecnica e letteratura scientifica, analizzando oltre 500 brevetti selezionati tra il 2001 e il 2018. La scelta di questo dataset è motivata dalla sua capacità di fornire una rappresentazione semantica dettagliata delle tecniche di difesa informatica, incluse le loro relazioni con tattiche offensive modellate nel framework MITRE ATT&CK.

D3FEND è stato scelto poiché rappresenta un’istanza reale del dominio di interesse, offrendo un’ampia copertura e specificità nel contesto della Cyber Situation Awareness, con il potenziale di supportare sistemi avanzati di RAG.

1. ***Descrizione del dimostratore che sarà realizzato nell’ambito della tesi***
   * *Laddove non sia prevista la realizzazione di alcun dimostratore, specificare il motivo per cui non è possibile realizzare un dimostratore da mostrare live (anche nella forma di un video) durante la seduta di laurea*

Sarà prodotto un video esplicativo che illustra il funzionamento del sistema. In particolare, verranno mostrate le risposte del sistema con e senza contesto.

**Proposta di sommario**

Nota: Deve contenere i titoli dei capitoli e delle sezioni. Questi capitoli sono obbligatori, ma è possibile integrare con capitoli aggiuntivi.

1. Introduzione
   * Definizione del problema
   * Rilevanza della problematica nel contesto dell’ingegneria informatica
2. Stato dell’arte
   * Analisi di dettaglio dello stato dell’arte
   * Retrieval Augmented Generation
   * I Knowledge Graph e la semantica
   * Individuazione di possibili avanzamenti rispetto allo stato dell’arte
3. Contributo originale alla soluzione del problema
   * Definizione della metodologia proposta
   * Progettazione del sistema proposto
   * Innovazioni di carattere applicativo
   * Strumenti, tecnologie e modelli utilizzati per la realizzazione
4. Validazione sperimentale e aspetti applicativi
   * Descrizione delle metriche di valutazione dei risultati
   * Definizione del protocollo sperimentale o di verifica
   * Descrizione dei dati e/o del caso di studi utilizzato per la sperimentazione
   * Presentazione e analisi dei risultati
   * Valutazione della significatività dei risultati ottenuti e dei miglioramenti apportabili

**Bibliografia**

[1] Sikos, L.F. Cybersecurity knowledge graphs. *Knowl Inf Syst* **65**, 3511–3531 (2023).

[2] Rahman, Moqsadur, et al. *Retrieval Augmented Generation for Robust Cyber Defense*. No. PNNL-36792. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), 2024.

[3] Kaloroumakis, Peter E., and Michael J. Smith. "Toward a knowledge graph of cybersecurity countermeasures." *The MITRE Corporation* 11 (2021): 2021.

[4] MITRE. "Common Vulnerabilities and Exposures (CVE)." MITRE Corporation. URL: [https://cve.mitre.org/](https://cve.mitre.org/" \t "_new).

[5] MITRE. "Common Weakness Enumeration (CWE)." MITRE Corporation. URL: <https://cwe.mitre.org/>.

[6] MITRE. "Common Attack Pattern Enumeration and Classification (CAPEC)." MITRE Corporation. URL: <https://capec.mitre.org/>.

[7] MITRE. "Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK)." MITRE Corporation. URL: <https://attack.mitre.org/>.

**Contributi di natura progettuale ed implementativa che il tesista dovrà fornire (circa 10 righe)**

Il tesista sarà responsabile di diversi aspetti progettuali e implementativi nell’ambito della tesi. Tra i contributi principali, si evidenzia:

* La progettazione e costruzione di un grafo di conoscenza rappresentativo del dominio CyberSA, basato su dataset specifici e tecnologie semantiche.
* L’uso di un modello di Knowledge Graph Embedding (KGE) per estrarre rappresentazioni vettoriali di entità e relazioni.
* L’integrazione di un sistema di Retrieval utilizzando modelli di similarità vettoriale, ottimizzato per query in linguaggio naturale.
* L’adattamento di un Large Language Model (LLM) per la fase di generazione di risposte basate su contesto, assicurando coerenza e rilevanza delle informazioni generate.
* Lo sviluppo di un ambiente di test per simulare scenari realistici di CyberSA e per valutare le performance complessive del sistema in condizioni operative.
* L’iterazione sul progetto per migliorare le prestazioni del sistema, attraverso test e ottimizzazioni specifiche.

**Tecnologie e materiali da impiegare durante l’attività di tesi**

* + *Dettagliare se il tesista avrà accesso a specifici strumenti (es: strumenti di misurazione, piattaforme robotiche, server messi a disposizione dal gruppo di ricerca, ad esempio equipaggiati con GPU)*

Durante l’attività di tesi, il tesista avrà accesso a diverse tecnologie per il completamento del progetto. Tra i principali:

* librerie python per il processamento e la gestione dei dati;
* strumenti specifici per la modellazione e l’elaborazione di grafi di conoscenza;
* un database vettoriale per la memorizzazione e il recupero di rappresentazioni vettoriali derivate da entità e relazioni del knowledge graph;
* LLM pre-addestrati per la gestione delle query e la generazione delle risposte;
* dataset specifici relativi al dominio della cybersecurity, utilizzati per creare la base di conoscenza.

**Eventuali esami ancora da sostenere alla data di consegna del presente documento**

* + Indicare per ciascun esame la data (mese/anno) in cui si prevede di sostenerlo

XXX

**Data (mese/anno) in cui l’attività di tesi è iniziata con un impegno sostanzialmente a tempo pieno**

Dicembre 2024

**Data (mese/anno) in cui presumibilmente sarà discussa la tesi di laurea**

Marzo/aprile 2025

***(Solo per le Tesi svolte in ERASMUS)***

**Sede**

XXX

**Periodo di svolgimento**

XXX

**Tutor**

XXX

Note