DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova – Dipartimento di Ingegneria dell'Informazione
Corso di Laurea in Ingegneria Informatica

Relazione per la prova finale
«Graph Anomaly Detection: Metodologie e Applicazioni»

Relatore: Prof.ssa Ornella Irrera

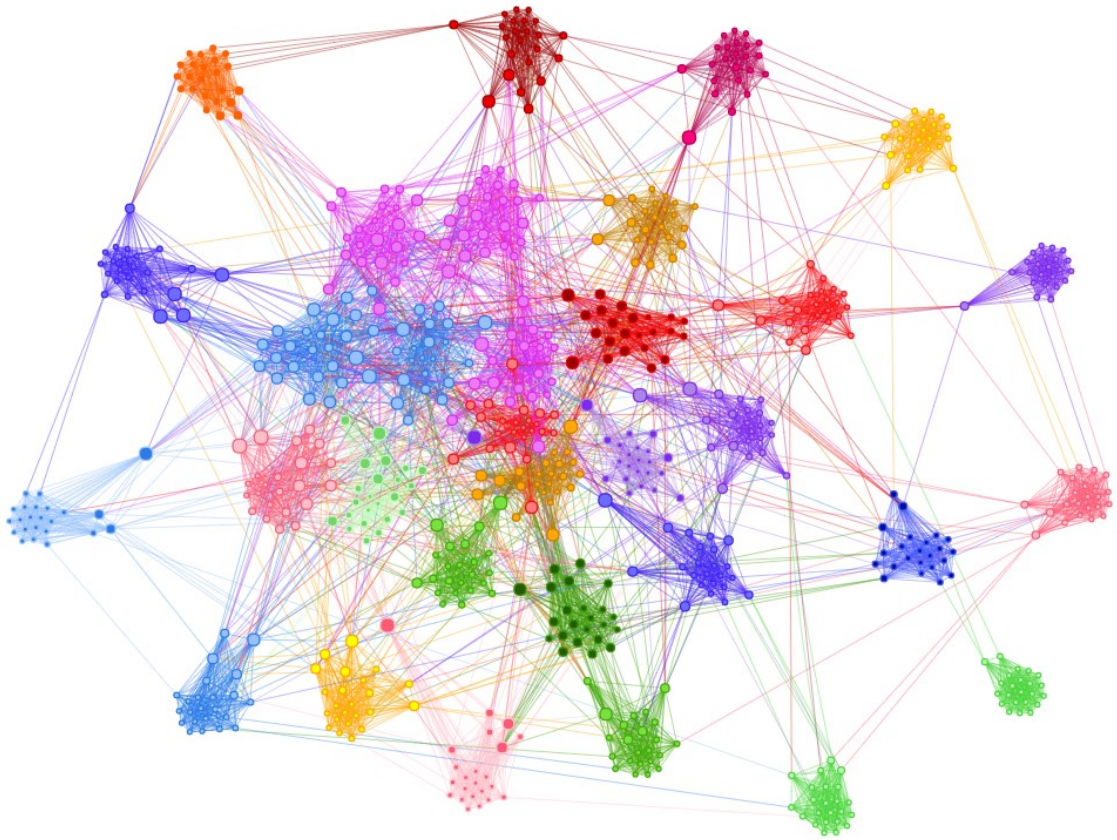
Laureando: Tiozzo Brasiola Alessandro

Matricola n.: 2101751

Padova, 22/09/2026.

INTRODUZIONE

- I dati reali sono sempre più caratterizzati da **relazioni complesse** tra entità
- I **grafi** consentono di rappresentare queste relazioni in modo naturale
- Le **anomalie** emergono spesso dalla **struttura** delle connessioni e non dai singoli elementi
- Identificarle è essenziale per comprendere il comportamento del sistema analizzato

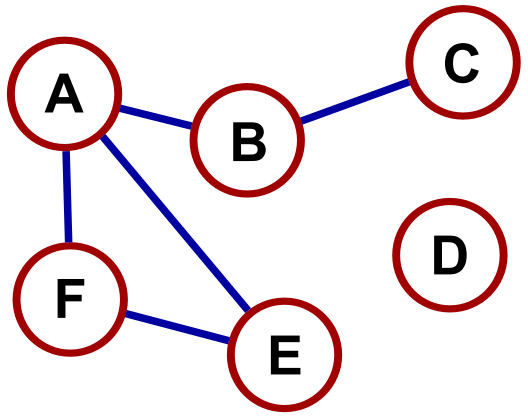


COS'È UN GRAFO

Un modello per rappresentare relazioni

- **Nodo** (Vertex)
- **Arco** (Edge)

$$G = (V, E)$$



Esempi

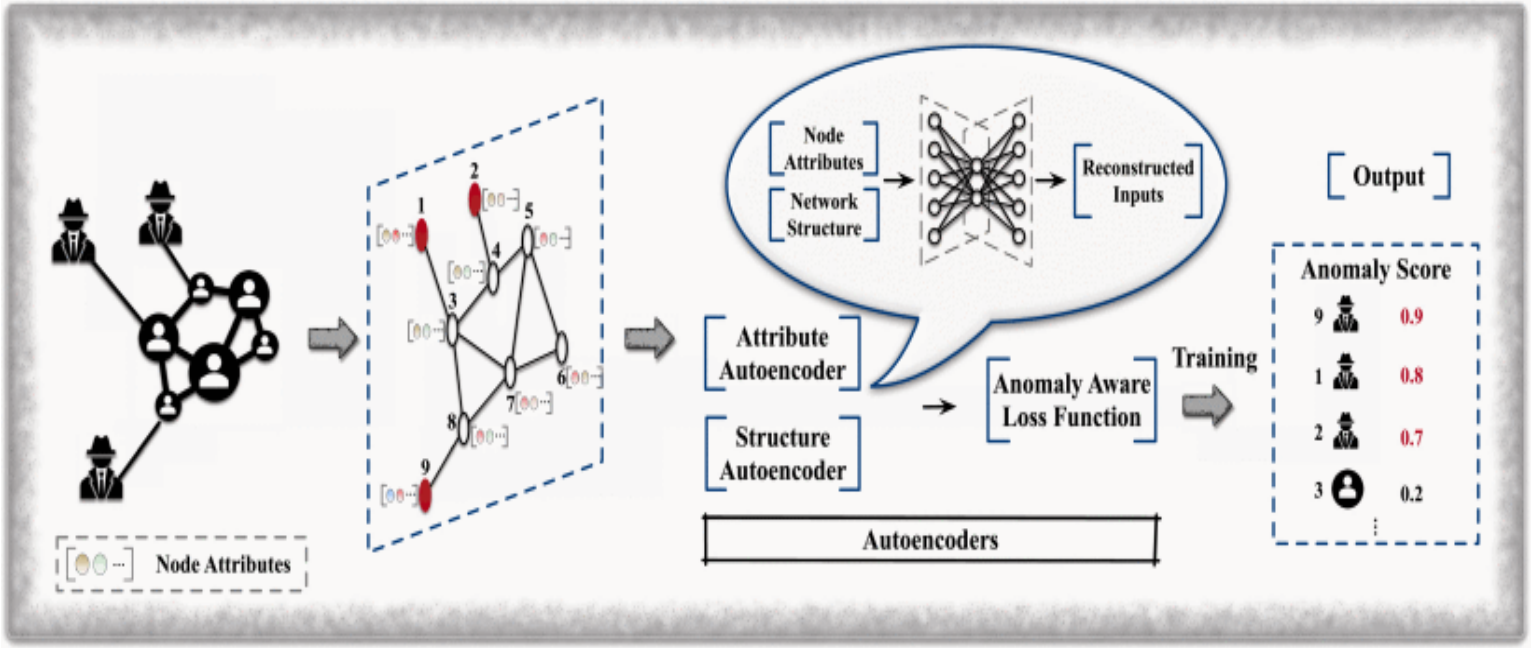
- Social network → utenti e amicizie
- Finanza → conti e transazioni
- Reti informatiche → dispositivi e connessioni

COS'È LA GRAPH ANOMALY DETECTION

Obiettivo: identificare nodi con caratteristiche o comportamenti differenti rispetto al contesto del grafo

Processo di rilevamento

- **Analisi** della struttura del grafo
- **Apprendimento** dei pattern normali
- **Identificazione** delle deviazioni
- Assegnazione di un **Anomaly Score**



TIPOLOGIE DI ANOMALIE

Le anomalie possono manifestarsi a diversi livelli del grafo

Node Anomalies

Nodi con comportamenti **insoliti**

Edge Anomalies

Relazioni **inaspettate** tra nodi

Subgraph Anomalies

Gruppi di nodi con **pattern anomali**

Graph Anomalies

Interi grafi **differenti** dal comportamento normale

TECNICHE PER LE ANOMALIE SUI NODI

Graph Autoencoder (GAE)

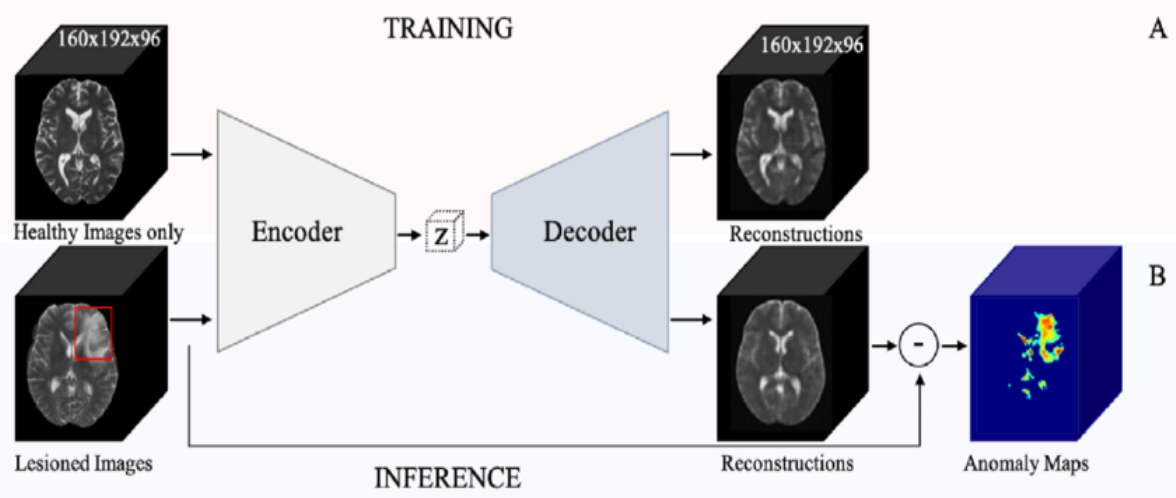
- **Comprimono** le informazioni del nodo e del suo vicinato in una rappresentazione latente
- Un elevato errore di ricostruzione può indicare un'anomalia

Contrastive Learning

- Apprende rappresentazioni **distinguendo** esempi simili e dissimili
- I nodi anomali tendono ad avere embedding differenti rispetto ai pattern normali

Graph Attention Networks (GAT)

- Attribuiscono pesi diversi ai nodi vicini
- Permettono di **individuare** nodi che si **discostano** dalle normali relazioni strutturali



TECNICHE PER LE ANOMALIE SU ARCHI E SUBGRAFI

Link Prediction

- Modella la **probabilità** di esistenza delle connessioni
- Archi poco probabili possono essere considerati anomali

Community Detection

- Individua gruppi **fortemente** connessi
- Subgrafi che si discostano dalla normale organizzazione possono essere anomali

Graph Neural Networks

- Analizzano le relazioni tra nodi collegati
- Evidenziano **connessioni inconsuete** rispetto alla struttura globale

Motif Analysis

- Analizza **pattern** ricorrenti di connessioni
- Configurazioni rare o inattese possono segnalare comportamenti anomali

TECNICHE PER LE ANOMALIE DINAMICHE

Temporal Graph Networks

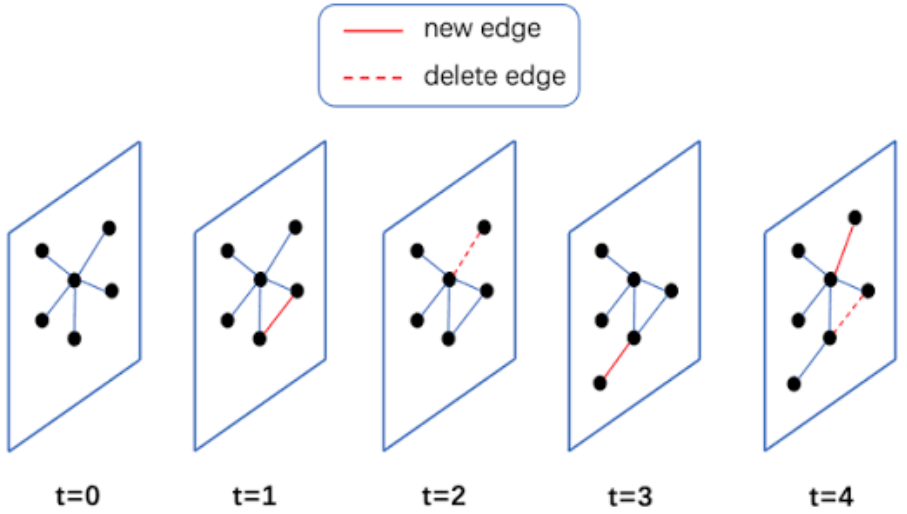
- Modellano l'**evoluzione** del grafo nel tempo
- Rilevano cambiamenti **improvvisi** nelle interazioni

Change Point Detection

- Identifica **variazioni significative** nella dinamica della rete
- Utile per intercettare attacchi, frodi o eventi anomali emergenti

Sequence Modeling (RNN, LSTM, Transformer)

- Analizzano sequenze temporali di eventi
- Permettono di individuare comportamenti inattesi rispetto alla **storia passata**



METRICHE DI VALUTAZIONE

Precision

- **Percentuale** di anomalie segnalate che sono realmente anomalie
- Riduce i falsi positivi

F1-Score

- **Media** armonica tra Precision e Recall
- Misura il **bilanciamento** complessivo del modello

Recall

- **Percentuale** di anomalie effettivamente individuate
- Riduce i falsi negativi

AUC-ROC e AUC-PR

- Valutano la **capacità discriminativa** del modello
- AUC-PR è particolarmente utile con dataset sbilanciati

DATASET DI RIFERIMENTO

Dataset utilizzati nella valutazione degli algoritmi

- YelpChi → recensioni fraudolente
- Amazon → spam e fake reviews
- Elliptic → frodi nelle transazioni
- TwitterBot-20 → rilevamento di bot
- CICIDS2017 → intrusion detection
- Cora → rete di citazioni scientifiche

DATASET	DOMINIO	NODI	ARCHI	% ANOMALIE
YELPCHI	Frodi recensioni	~45K	~3.6M	14.5%
AMAZON	Frodi recensioni	~11K	~4.4M	6.87%
ELLIPTIC	Transazioni Bitcoin	~200K	~234K	~2%
TWITTERBOT-20	Bot detection	~229K	~33M	~5%
CICIDS2017	Intrusioni rete	~2.8M	variabile	~20%
CORA	Citation network	~2.7K	~5.4K	iniettate

APPLICAZIONI

Cybersecurity

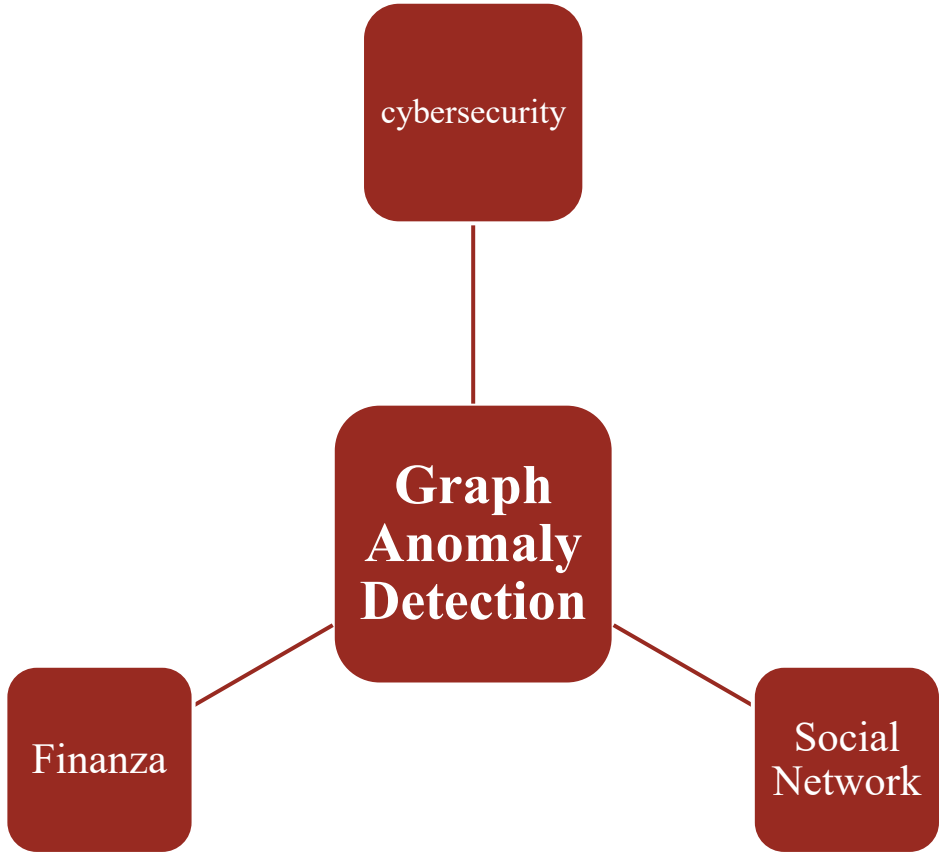
- Connessioni sospette tra dispositivi
- Intrusioni e attacchi di rete

Finanza

- Rilevamento di frodi nelle transizioni
- Riciclaggio di denaro

Social Network

- Bot e account falsi
- Campagne coordinate di disinformazione



CONCLUSIONI

Principali evidenze

- I grafi permettono di modellare efficacemente dati relazionali complessi
- Le anomalie possono emergere dalla struttura delle connessioni
- Le tecniche basate su **Deep Learning** e **GNN** sono oggi le più efficaci

Sfide future

- Scalabilità su grafi di **grandi dimensioni**
- **Interpretabilità** dei risultati
- **Gestione** di grafi dinamici ed eterogenei