COMPUTER AND CONTROL ENGINEERING

DAUIN/Centro SmartData - Exploring the Social Media and Instant Messaging Networks and Ecosystems: Implications for Cybersecurity and Misinformation

Funded By	Dipartimento DAUIN Centro Interdipartimentale SmartData@PoliTO

Supervisor	MELLIA MARCO - marco.mellia@polito.it

Contact	

Context of the research activity	Social media (such as Facebook, LinkedIn) and instant messaging applications (such as Telegram and Whatsapp) are platforms that connect people, thanks to groups and channels, creating time-evolving networks and graphs. These networks carry information on users' behavior and the reuse of media content. The research aims to shed light on these complex ecosystems, discover interesting patterns (e.g., polarization) and highlight possible misuse, from cybersecurity (e.g., phishing) to misinformation.

Objectives	Research objectives: The expected outcome of the research would be a deeper understanding of the dynamics of social media and instant messaging networks, including identifying key patterns, behaviors, and cyber-threats. The research will follow the following steps: - Data collection process, involving web crawlers to gather data from social media, and novel solution to automatically crawl instant messaging platforms. - Network creation and study the characteristics of these networks, both in considering a single snapshot, and how these change over time. - Understanding of how users interact with and within these networks, and what behavior patterns can be observed. - Highlight the potential anomalies and security threats of these platforms and how they can be detected and addressed using automatic approaches. - Understanding how these platforms can be used to spread misinformation and inform strategies for countering it. - Inform strategies for countering cyber threats and misinformation. The analysis will be performed with machine learning techniques, graph analysis, and natural language processing tools. Outline of the research work plan 1st year - Study of the state-of-the-art of social network and instant messaging data-driven data collection and analysis - Design and deployment of a data collection process with automatic and

scalable crawlers for the web, social networks, and instant messaging platforms.
- The data will be stored, organized, and analyzed using big data techniques (such as Pyspark).

2nd year
- Adaptation and extension of existing solutions to study the networks
- Propose and develop innovative solutions to the problems of cyber threats and misinformation

3rd year
- Tune the developed techniques and highlight possible strategies to counteract the various threats
- Application of the strategies to new data for validation.

List of possible venues for publications
- Elsevier Online Social Networks and Media
- ACM Transactions on the Web
- Springer Social Network Analysis and Mining
- IEEE/ACM ASONAM
- ACM WebSci

Skills and competencies for the development of the activity	- Programming skills (python, big data) - Machine learning - Data science for network analysis - Basics of parsing and scraping/crawling web pages