

# Team Meeting 2018

Saharnaz Dilmaghani



B.Sc Information Technology  
University of Tabriz, Iran  
2010 – 2014



PhD candidate at PCOG  
SnT, Luxembourg  
Since March 2018



M.Sc Computer engineering  
Bilkent, Turkey  
2015 – 2018

# Background & Interests

Privacy

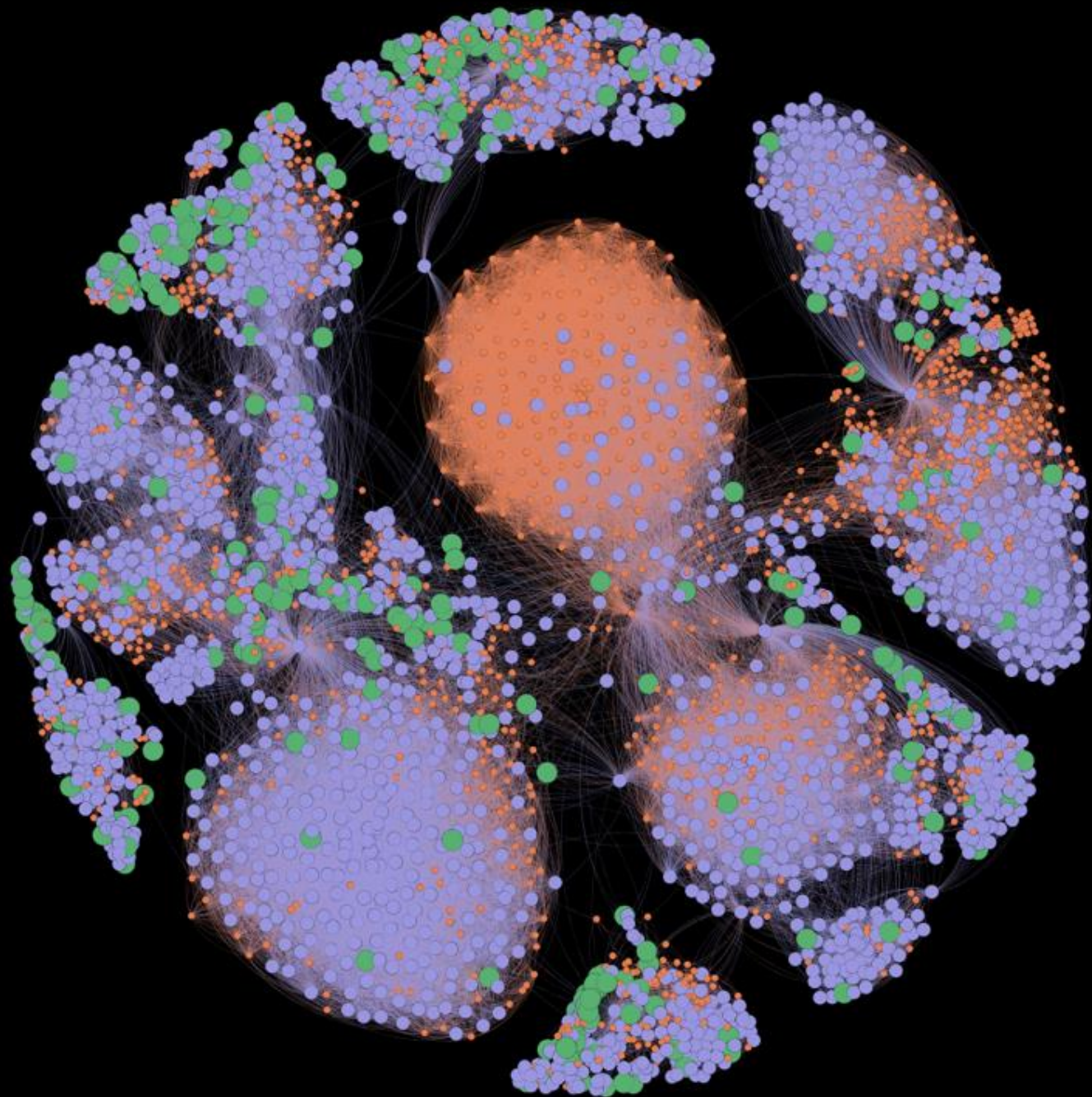
Big data analysis

Network science

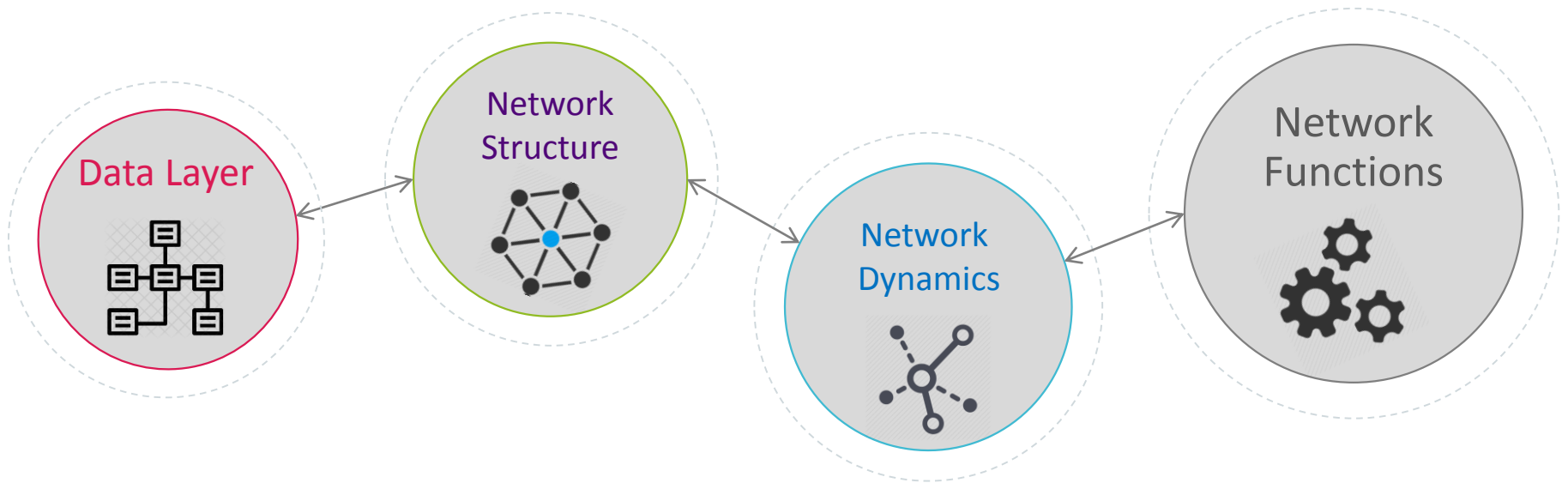
PhD topic

“Adaptive Clustering Algorithms”









### Data Layer

Translating relational  
data to network

Definition of nodes and  
edges

### Network Structure

Similarity functions  
Network topologies

### Network Dynamics

Similarity functions  
i.e. KHOPCA, LPA

### Network Functions

i.e. brain networks

## Network-based Machine Learning for Privacy Preservation in Bio-medical Big Data Analytics

Saharnaz E. Dilmaghani<sup>1</sup>, Matthias R. Brust<sup>1</sup>, Johnatan Pecero<sup>3</sup>,  
Grégoire Danoy<sup>2</sup>, Pascal Bouvry<sup>1,2</sup>

<sup>1</sup>Interdisciplinary Centre for Security Reliability and Trust (SnT), University of Luxembourg

<sup>2</sup>Computer Science and Communications Research Unit, University of Luxembourg

<sup>3</sup>Agence pour la normalisation et l'économie de la connaissance (ANEC), Luxembourg

Machine learning promises to bring valuable insights needed for innovation on the growing volume and complexity of data. Particularly critical big data such as bio-medical data, which is generated continuously and from diverse sources (e.g. smart watches and medical records) requires not only optimal results but also the guaranteed anonymization of the individual's data. As it turns out, existing machine learning techniques are unable to cope with these emerging challenges of big data. A widely used process for machine learning is clustering that brings valuable inference of data such as medical diagnosis for the bio-medical data.

In this study, we focus on using network structures to design an adaptable and localized clustering algorithm that provides existing machine learning techniques with privacy preserving feature. Although omni-present in data, networks are often discarded in machine learning. Therefore, we propose using a network-based approach to benefit from network science to transform big data into network structures, which enables to dynamically cluster data while preserving privacy. The method applies a set of defined rules (based on a probability similarity function) in order to generate the nodes and connectivity patterns of the network from the source data. The links are generated such that stronger connections through the nodes reflect the likelihood of belonging to the same cluster. In addition, while generating the network, the  $k$ -degree anonymity technique is applied to decrease possible correlations between sensitive data which results in an enhanced privacy preservation machine learning technique. Benefiting from the privacy exposure metric [2], we measure the privacy quality of the proposed scheme. The network structure is beneficial for providing not only privacy preservation but also adaptability on the scheme.

We measure the performance of our approach on bio-medical big data for Parkinson's Disease Map [3]. Our approach provides privacy preservation to bio-medical data while improving the learning of diagnosis by adding new data regarding a certain disease. Additionally, the proposed technique is adaptive to dynamic data and has low computational complexity. We expect that our algorithm has an impact on other applications of big data such as time series dataset. Moreover, we also plan to provide experiments on comparing the clustering quality of our method with existing machine learning techniques.

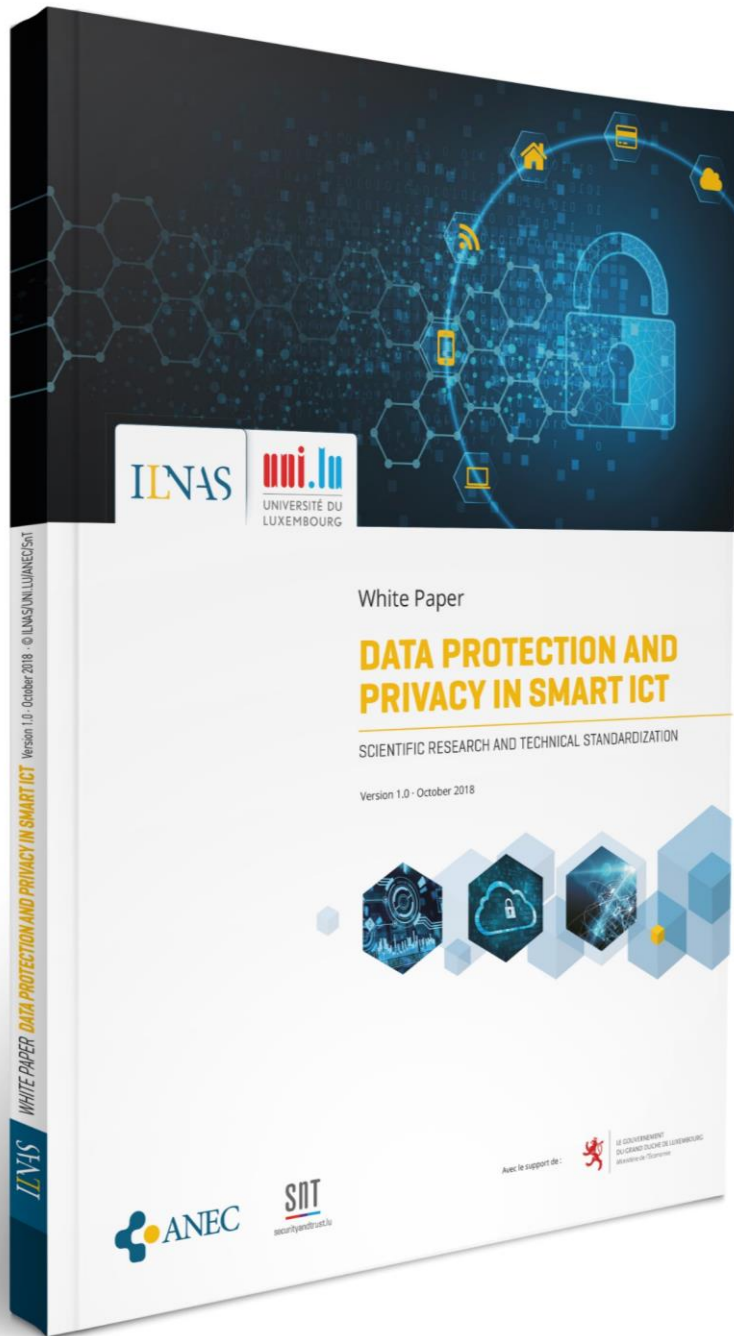
### Acknowledgment

This work partially funded by the joint research programme UL/SnT-ILNAS on Digital Trust for Smart-ICT.

### References

- [1] M. R. Brust, H. Frey, and S. Rothkugel, "Adaptive multi-hop clustering in mobile networks," in Proc. of Inter. Conference on mobile technology, applications, pp. 1327138, ACM, 2007.
- [2] Fang-Jing Wu, M. R. Brust, et al. "The privacy exposure problem in mobile location-based services." Global Communications Conference (GLOBECOM), 2016 IEEE. IEEE, 2016.
- [3] LCSB-Luxembourg Centre for Systems Biomedicine, Available at <https://wwwfr.uni.lu/lcsb/>

- Network structure for machine learning
  - Adaptive
  - Dynamic
- June 2018
- NetSci 2018



- Privacy and data protection
  - Big data
  - Cloud computing
  - IoT
- Scientific solutions
- Standardization
- October 2018

# Transforming Collaboration Data into Network Layers for Enhanced Analytics

Saharnaz E. Dilmaghani<sup>1</sup>, Apivadee Piyatunrong<sup>2</sup>, Pascal Bouvry<sup>3</sup>, and Matthias R. Brust<sup>1</sup>

<sup>1</sup> Interdisciplinary Centre for Security, Reliability, and Trust (SnT)  
University of Luxembourg  
`firstname.lastname@uni.lu`

<sup>2</sup> National Electronics and Computer Technology Center (NECTEC)  
A member of NSTDA, Thailand  
`apivadee.piy@nectec.or.th`

<sup>3</sup> Faculty of Science, Technology and Communication (FSTC)  
University of Luxembourg  
`pascal.bouvry@uni.lu`

**Abstract.** We consider the problem of automatically generating networks from data of collaborating researchers. The objective is to apply network analysis on the resulting network layers to reveal supplemental patterns and insights of the research collaborations. In this paper, we describe our data-to-networks method, which automatically generates a set of logical network layers from the relational input data using a linkage threshold. We, then, use a series of network metrics to analyze the impact of the linkage threshold on the individual network layers. Moreover, results from the network analysis also provide beneficial information to improve the network visualization. We demonstrate the feasibility and impact of our approach using real-world collaboration data. We discuss how the produced network layers can reveal insights and patterns to direct the data analytics more intelligently.

**Keywords:** Collaboration data · data-to-network · network visualization · data analysis · interaction methods

- Collaboration data
  - NECTEC
- Define a Linkage Threshold to convert data to network layers
- Network metrics
- November 2018
- OLA 2019



# THANKS!



You can find me at  
[saharnaz.dilmaghani@uni.lu](mailto:saharnaz.dilmaghani@uni.lu)