

Post-doctoral position (18 to 24 months) at INRAE - UMR TETIS, Montpellier, France

Topic

Spatio-Temporal analysis of epidemiological events on complex networks issued from large volumes of heterogeneous data

Keywords

Complex networks, network analysis, multi-scale, spatio-temporal resolution, heterogeneous data

Context

In recent years, the amount of data generated on human and animal health events has increased significantly. Epidemiologists must therefore regularly analyze these data with various spatial and temporal resolutions. The proposed postdoc contract is part of the H2020 MOOD project "Monitoring Outbreak events for Disease surveillance in a data science context" (<https://mood-h2020.eu>), which brings together 25 partners from 10 countries. This project is led by CIRAD (UMR ASTRE) and aims at improving the detection, monitoring and evaluation of emerging infectious diseases in Europe by using advanced data science techniques on massive multisource data.

The work package 3 "data ingestion and integration" is centered on the linking of heterogeneous data collected and processed in the context of the MOOD project. These data are heterogeneous in terms of domain (e.g., medical, environmental, social) and in terms of format (e.g., textual data, satellite imagery, multivariate quantitative data), and can be originated by both official (e.g., medical institutes, scientific laboratories) and unofficial (e.g., newspapers, social media) sources. By consequence, this diversity is also reflected in the spatial and temporal scales of the data.

More precisely, in the context of this post-doc, we are interested in modeling information about epidemiological events (detected from various data sources that are syntactically and semantically heterogeneous) into complex networks models that can allow advanced spatio-temporal analyses.

Methodology

The postdoc is focused on the possibility to model the heterogeneous data collected and processed in the context of the MOOD project into advanced complex network models, i.e., networks that integrate spatial and temporal information about the data.

The objective is twofold: (i) to show how heterogeneous data about an epidemiological event can be integrated, aggregated and analyzed into complex network models in order to allow an analysis of the complex spatio-temporal phenomena that characterize the life cycle of an epidemic, and (ii) to define original networks analysis and data science techniques in order fully exploit the information modeled in such spatio-temporal networks.

The research question at the center of this postdoc can be formulated as follows: How can we relate

spatio-temporal information from epidemic-related data in order to have a spatio-temporal analysis framework in the One Health context?

More precisely, we wish to propose generic methods to link and aggregate information from heterogeneous sources (in particular official and unofficial sources) into feature-rich networks able to embed spatio-temporal features, that will allow to analyze the life cycle of an epidemic according to its spatial and temporal evolution.

The final aim is then to bring new knowledge to experts, that will represent a precious complement to the classic source of information already exploited in the project. This spatio-temporal linking process will have to take into account some reliability and quality factors associated with the different descriptors, i.e., depending on source types and on the confidence of the algorithms in use.

Gross Salary

2300 to 2900 based on previous professional experience.

Candidate profile

PhD in computer science.

Preference will be given to highly motivated candidates with research experience in complex network analysis, heterogeneous data science and data science applied to epidemiology related tasks.

Application instructions:

Qualified applicants are invited to send their application to Maguelonne Tisseire (maguelonne.teisseire@inrae.fr) and Roberto Interdonato (roberto.interdonato@cirad.fr) as a single pdf file containing a cover letter describing their research background and motivation, a detailed CV and the contact details of up to three referees.

Application deadline: June 28, 2021

Interviews for selected candidates : July 2, 2021

Bibliography

R. Adderley, P. Seidler, A. Badii, M. Tiemann, F. Neri, M. Raffaelli. Semantic Mining and Analysis of Heterogeneous Data for Novel Intelligence Insights. Proc. of The Fourth International Conference on Advances in Information Mining and Management, IARIA, p.36-40, 2014

Goel R., Sallaberry A., Fadloun S., Roche M., Valentin S., Poncelet P. EpidNews: An epidemiological news explorer for monitoring animal diseases. In : Proceedings of the 11th International Symposium on Visual Information Communication and Interaction (VINCI 2018), Växjö, Suède, Août 2018.

P. Cimiano, L. Schmidt-Thieme, A. Pivk, S. Staab. Buchtitel: Learning Taxonomic Relations from Heterogeneous Evidence. Proc. of the ECAI 2004 Ontology Learning and Population Workshop, 2004

A. Henriksson, J. Zhao, H. Boström, H. Dalianis. Modeling Heterogeneous Clinical Sequence Data in Semantic Space for Adverse Drug Event Detection. Proc. of IEEE Int. Conf. on Data Science and Adv. Analytics, 2015

Roberto Interdonato, Raffaele Gaetano, Danny Lo Seen, Mathieu Roche, Giuseppe Scarpa: Extracting multilayer networks from Sentinel-2 satellite image time series. Netw. Sci. 8(S1): S26-S42 (2020)

Roberto Interdonato, Matteo Magnani, Diego Perna, Andrea Tagarelli, Davide Vega: Multilayer network simplification: Approaches, models and methods. Comput. Sci. Rev. 36: 100246 (2020)

Roberto Interdonato, Martin Atzmueller, Sabrina Gaito, Rushed Kanawati, Christine Largeron, Alessandra Sala: Feature-rich networks: going beyond complex network topologies. Appl. Netw. Sci. 4(1): 4:1-4:13 (2019)

Matteo Magnani, Obaida Hanteer, Roberto Interdonato, Luca Rossi, and Andrea Tagarelli. Community detection in multiplex networks. CoRR, abs/1910.07646, 2021. (*to appear on ACM Computing Surveys*)