



HAL
open science

Inferring the transmission dynamics of Avian Influenza from news and environmental data

Nejat Arinik, Roberto Interdonato, Mathieu Roche, Maguelonne Teisseire

► To cite this version:

Nejat Arinik, Roberto Interdonato, Mathieu Roche, Maguelonne Teisseire. Inferring the transmission dynamics of Avian Influenza from news and environmental data. International Conference on Complex Networks and their Applications, Nov 2022, Palerme, Italy. hal-03936660

HAL Id: hal-03936660

<https://hal.science/hal-03936660>

Submitted on 12 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inferring the transmission dynamics of Avian Influenza from news and environmental data

Nejat Arinik¹, Roberto Interdonato², Mathieu Roche², and Maguelonne Teisseire¹

¹ INRAE - UMR TETIS, 500, Rue Jean François Breton, 34000 Montpellier France

² CIRAD - UMR TETIS, 500, Rue Jean François Breton, 34000 Montpellier, France,

`firstname.lastname@inrae.fr`¹,

`firstname.lastname@cirad.fr`²

1 Introduction

Avian Influenza (AI) is a highly contagious animal disease. Although a few human infection cases occurred in the past, it mainly infects wild and domestic bird species. Transmission between birds can be direct due to close contact, or indirect through contaminated materials such as feed and water. Particularly, migratory wild birds play a key role in this transmission and make the viruses spread over long distances. The emergence and spread of AI has serious consequences for animal health and a substantial socio-economic impact for worldwide poultry producers. Due to its highly contagious nature, it is critical to monitor the evolution and spread of this disease.

There are mainly two types of surveillance systems established for this purpose: Indicator-Based Surveillance (IBS) and Event-Based Surveillance (EBS). IBS usually relies on official notification procedures submitted by a country, whereas EBS extracts epidemiological events mostly from unofficial sources, such as news articles, through various Natural Language Processing tasks. Recently, several EBS platforms have shown their effectiveness by detecting the first signals of emerging infectious disease outbreaks in a timely manner and providing alerts within previously unaffected areas (e.g. [4]).

Even though it is a well-studied and constantly monitored disease, real transmission routes of the AI viruses are hard to discover, i.e. where did an outbreak that appears in some location come from. Nonetheless, acquiring such disease transmission routes may allow to discover transmission patterns and to timely predict subsequent outbreaks. In this work, we focus on the problem of inferring how the AI disease spreads in the absence of knowledge about real disease transmission routes. We address it as a network inference problem, leveraging on spatio-temporal event datasets provided by EBS platforms and environmental data. The inferred network is then used to study the transmission patterns of the AI disease. Although network inference problems have been widely studied in the domain of information diffusion, our work differentiates from others in that 1) unlike the contact network analysis which are at the individual farm level, it is at the metapopulation level (e.g., country, province, city) [5], and 2) it reflects a more realistic scenario by taking advantage of the migration routes of wild birds, as well as poultry transportation data and environmental data related to the outbreak locations.

In this ongoing work, our aim is to illustrate the relevance of our proposed method by applying it to the AI outbreak datasets collected by an EBS platform, and show

how it can be put in practice. Here, we summarize the proposed method and give some preliminary results.

2 Method

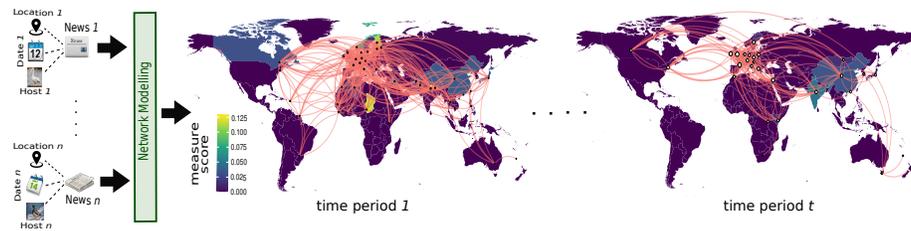


Fig. 1. Network modelling for inferring the transmission dynamics of Avian Influenza from news and environmental data.

The method we propose is designed to study the transmission patterns of the AI disease, and it is constituted of two parts. The first part aims at tackling the network inference problem introduced in Section 1 and illustrated in Figure 1. To construct a disease transmission network, we rely on a sequential probability based approach and reconstruct the transmission routes based on how AI outbreaks evolve in the news articles collected by EBS platforms, accompanied by migration routes of wild birds, poultry transportation data and environmental data related to the outbreak locations. In the end, we obtain an attributed location-aware dynamic network.

We define this dynamic network as a sequence of time ordered static networks, where each static network encodes the aggregated outbreak information for a particular time period (e.g. season). The nodes of the network correspond to geographical locations (e.g., country, province, city) extracted in news data and provided by surveillance platforms [4], and the edges between them represent the disease transmission routes that are more likely to occur. We determine these transmission routes and the associated probabilities in three steps. First, we identify pairs of locations that are connected through disease transmission routes. At two adjacent time intervals t and $t + 1$, all the locations that have infection cases at time t are connected to those locations that have infection cases at time $t + 1$. Then, we calculate the transmission probabilities based on three aspects: 1) temporal differences between outbreaks and distance between outbreak locations, 2) risk zones defined based on environmental data, and 3) poultry transportation and migratory routes of wild birds. These transmission probabilities determine the edge weights in our network. Finally, we filter out edges whose weight is lower than a given threshold.

In the second part, we study the resulting dynamic network leveraging on two groups of state-of-the-art spatio-temporal measures. The first group characterizes a

single node through macroscopic measures, which allow identifying key locations in disease transmissions. We consider the temporal betweenness centrality, the adapted Pagerank centrality [3], the super-spreader and super-susceptible indices [1]. The second group of measures aims to summarize the whole network structure. We consider the spatial Gini coefficient and the time-space spreading index based on hotspot analysis [3]).

3 Preliminary results

To show the interest of our method, we apply it to the AI event datasets collected by PADI-Web [4], a well-known surveillance platform in Epidemic Intelligence, for the period of 2020-2021. We first empirically evaluate our inferred AI transmission network based on the phylogenetic analysis tool *NextStrain* [2] and the transmission probability results published in [6] for several strains of AI (e.g. H5N1, H5N8) from the beginning of 2020 to the end of 2021. Although our network is three times denser than the result of the phylogenetic analysis, we correctly inferred almost half of the AI transmission routes across countries. Nonetheless, it is worth noting that since the phylogenetic analysis is used to reveal virus genetic changes in space and time, it could be a solution to the network validation task only at a coarser level.

Our preliminary results also confirm the existence of hotspots and super-spreaders, i.e. the locations being particularly effective in transmitting AI disease, in our dataset. We observe that the identified hotspots are spatially concentrated mostly in Asia in the beginning of 2020, then the hotspots spread out towards Europe and North America in the subsequent time periods. Furthermore, China and India appear as the main super-spreaders in the beginning of 2020, then Russia dominates these countries starting from the winter season of 2021.

References

- [1] W. C. B. Chin and R. Bouffanais. “Spatial super-spreaders and super-susceptibles in human movement networks”. In: *Scientific Reports* 10.1 (Oct. 2020).
- [2] J. Hadfield, C. Megill, S. M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, and R. A. Neher. “Nextstrain: real-time tracking of pathogen evolution”. In: *Bioinformatics* 34.23 (May 2018). Ed. by Janet Kelso, pp. 4121–4123. DOI: 10.1093/bioinformatics/bty407.
- [3] M. Nanni, L. Tortosa, J. F. Vicent, and G. Yeghikyan. “Ranking places in attributed temporal urban mobility networks”. In: *PLOS ONE* 15.10 (2020). Ed. by Alireza Abbasi, e0239319.
- [4] S. Valentin, E. Arsevska, J. Rabatel, S. Falala, A. Mercier, R. Lancelot, and M. Roche. “PADI-web 3.0: A new framework for extracting and disseminating fine-grained information from the news for animal disease surveillance”. In: *One Health* 13 (2021), p. 100357.
- [5] X. Yang, J. Liu, X.-N. Zhou, and W. K. W. Cheung. “Inferring disease transmission networks at a metapopulation level”. In: *Health Information Science and Systems* 2.1 (2014).
- [6] J. Zhang, X. Li, X. Wang, H. Ye, B. Li, Y. Chen, J. Chen, T. Zhang, Z. Qiu, H. Li, W. Jia, M. Liao, and W. Qi. “Genomic evolution, transmission dynamics, and pathogenicity of avian influenza A (H5N8) viruses emerging in China, 2020”. In: *Virus Evolution* 7.1 (Jan. 2021).