



ERCIM "ALAIN BENSOUSSAN"  
FELLOWSHIP PROGRAMME



## Scientific Report

First name / Family name	Panteleimon Vikatos
Nationality	Greek
Name of the <i>Host Organisation</i>	CNR IIT
First Name / family name of the <i>Scientific Coordinator</i>	Andrea Passarella
Period of the fellowship	01-03-2018 to 31/10/2018

### I – SCIENTIFIC ACTIVITY DURING YOUR FELLOWSHIP

Migration has gained the current interest of demographers and social scientists. Projecting migration rates and unveiling relationships are crucial. However, the limitation of available data has been one of the main bottlenecks for empirical analyses and for theoretical advances in the study of migrations. In particular, data about international migration flows are largely inconsistent across countries, typically outdated, and often non-existent.

Current studies declare that social media constitute a reliable source of information on migration flows. People constantly use online social networking sites sharing content about their daily lives and things that happen around them. These systems have revolutionized the way we communicate, by organizing our offline social relationships in a digital form. This massive popularity of online social media provides the opportunity to detect useful characteristics and patterns of their behaviour.

We were interested in extracting migration patterns on a large scale based on users' posts. The change of geolocated public streams during time intervals can depict mobility and migration flows. We consider that migration might influence the user's communication and connections in OSN environments. We focused to discover migration patterns and correlate migration phenomenon with potential changes in users' ego network in structure and communication frequency.

In addition, we interested to understand migration patterns of scientists through the analysis of research collaboration networks via Google Scholar/Elsevier Scopus crawling.

Our methodology for collecting the necessary information and discover the users that are categorized as migrants is described in the following independent tasks:

- Crawling. We gather geolocated tweets from the Twitter platform and affiliation history from academic publications database.
- User Characterization. The appropriate information of each user timeline is gathered. The recognition of user's migration relies on the change of his/her modal country.
- Analysis of ego-network. The post-processing analysis of the ego networks of each user during time intervals.

To identify active Twitter users we gathered data from the 1% random sample of all tweets, provided by the Twitter Stream API. We gather only the subsample of geolocated tweets. We crawl the recent 3,200 tweets of the obtained users. Also, the crawled users constitute the initial seed for a BFS procedure which is handled by a dedicated crawler that allows sampling of the Twitter data in this manner. The proposed crawler performs a Breadth-First search (BFS) starting from one (or multiple nodes) in order to traverse into Twitter. We store the crawled users using a unique identifier for each user account. Our collected dataset contains a large set of Twitter profiles including a sample of their public posts stream (tweets) and optionally the lists of user ids that correspond to friends/followers.

In terms of Elsevier Scopus database, we collected publications from seeds and we retrieved the affiliation history for each author of the papers.

The characterization of a Twitter user as a migrant is the modification of his/her modal country. We declared as Modal the country of residence for the given period from where most of the geolocated tweets were posted. If the uncertainty is large, i.e. if the number of tweets in the modal country is not at least three times as large as the number of tweets for the second most frequent country, then we discarded the information for that user for the specific period. We declared as Migrant a user who the modal country for two consecutive periods is not the same. For instance, for the first period of four months the modal country is A and then, for the second period, the modal country is B, we estimate that the user moved from country A to country B over the eight months considered.

In academic networks, we recognized a migrant through his/her affiliation history in other words if affiliation is changing a migration pattern is generating.

Our main goal was the analysis of ego network for migrants group. An ego network is formed of a single individual (ego) and the other users directly connected to it (alters). Ego networks are the graph-based abstraction that is used to study the social relations between an individual and alters. The ego-alter tie strength is typically computed as a function of the frequency of interactions between the ego and the alter.

## **II – PUBLICATION(S) DURING YOUR FELLOWSHIP**

--

## **III – ATTENDED SEMINARS, WORKSHOPS, CONFERENCES**

I attended to international conference 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) in Barcelona Catalonia.

## **IV – RESEARCH EXCHANGE PROGRAMME (REP)**

I had the opportunity to visit Telenor AI group at NTNU, Trondheim, Norway using the REP and to collaborate with assistant professor Heri Ramampiaro. During my visit, I presented my current research in members of Telenor AI group.