Metrics of Evolving Ego-Networks with Forgetting Factor

Rui Portocarrero Sarmento, University of Porto, Faculty of Engineering, Porto, Portugal

ABSTRACT

Nowadays, treating the data as a continuous real-time flux is an exigence explained by the need for immediate response to events in daily life. We study the data like an ongoing data stream and represent it by streaming egocentric networks (Ego-Networks) of the particular nodes under study. We use a non-standard node forgetting factor in the representation of the network data stream, as previously introduced in the related literature. This way the representation is sensible to recent events in users’ networks and less sensible for the past node events. We study this method with large scale Ego-Networks taken from telecommunications social networks with power law distribution. We aim to compare and analysis some reference Ego-Networks metrics, and their variation with or without forgetting factor.

KEYWORDS

Data Stream Analysis, Ego-Networks, Real-Time Applications, Social Network Stream Mining, Telecommunication Networks

INTRODUCTION

Streaming from a large network is known to be a hard problem to solve with typical hardware or software. It is common to see software and hardware present memory or processing limitations when doing the output of networks with more than a few thousand nodes and edges.

The authors discuss and propose the data input as a network data stream enabling the analysis with a commodity machine. Our contributions with this document are the description of a method for large scale online ego-network streaming focusing on the node level of the telecommunications network in the study. The proposed ego-network streaming is high efficient because of used and developed algorithms for analysis. The authors empirically demonstrate this streaming method can be used to represent ego-networks. There are several possible applications and large usability of ego-networks with node forgetting factor on the telecommunications network dataset as the authors demonstrate further in this document. The results were obtained by simulation of data streaming from databases.

This document is organized in the following manner:

The authors present related work, regarding ego-networks and their analysis metrics, in section Related Work. Then, in section Streaming Algorithms, the authors describe the developed algorithms for the forgetting factor application. After presenting the algorithms, the authors introduce the case study and the studied data in section Case Study. Section Case Study also contains the results of applying the algorithms to the described and selected data. Finally, in section Conclusions, the authors end this document with conclusions, discussion, and future work.

DOI: 10.4018/IJSODIT.2017010103

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
RELATED WORK

Several studies were presented regarding ego-networks in the scope of social network analysis. The areas are varied between biology, sociology to criminal networks. In this section, the authors introduce related work in a wider scope by mentioning publications regarding social networks, which include information about ego-networks.

Ego-Networks

In (Hanneman & Riddle, 2005), a throughout exposition about social networks is made, and a full chapter is dedicated to ego-networks. Hanneman et al. define “ego” as an individual “focal” node in a network. “neighborhood” sets the boundaries of ego networks and includes all the direct connections and egos that tie with an ego. Dejordy et al. (Dejordy & Halgin, 2009), introduce the network perspective and the differences between socio-centric and ego-centric analysis. The ego-centric approach fits studies about phenomena or entities across different networks. The socio-centric approach is more suitable for studying interaction within a defined network.

Wasserman et al. provide a complete study of social networks with several models in (Wasserman & Faust, 1994). Some relevant studies address the social structure of competition.

For Burt et al. (Burt, 1992), the social structure of competition addresses the consequences of voids in relational and resource networks. Competitive behavior can be understood regarding player access to “holes” in the social structure of the competitive arena. Those “structural holes” create entrepreneurial opportunities for information access, timing, referrals, and control. Ego-networks analysis provides an answer to this sensible information or properties that are also studied in the Case Study section of this document.

Figure 1 represents two ego-networks with connections to the 4th and 2nd order, respectively for node 1 and 5. This example shows how the same network change in terms of visualization, depending on the selected ego node.

Metrics

In this section, the authors introduce several metrics associated with ego-networks studies. The authors choose effective size, efficiency, and constraint to study the behavior of their developed node forgetting factor algorithm on the evolving streaming networks. First, let us proceed with a brief description of these metrics.

Hanneman and Riddle (Hanneman & Riddle, 2005), describe the metric effective size of the network as the number of alters that ego has, minus the average number of ties that each alter has to other alters. The authors give a simple example where a network node a has ties to three other actors. They continue and detail that none of these three has ties to any of the others. Hanneman and Riddle conclude effective size of ego’s network is three. Alternatively, suppose that a has ties to three others and that all of the others are linked to one another. Node a’s network size is three, but the ties are “redundant” because a can reach all three neighbors by reaching any one of them. The average degree of the others, in this case, is 2 (each alters tied to two other alters). So, the effective size of the network is its actual size (3), reduced by its redundancy (2), to yield an effective size of 1.

In Hanneman & Riddle, (2005), Hanneman and Riddle continue and describe efficiency, a normalization of the effective size of ego’s network by its actual size. That is, what proportion of ego’s ties to its neighborhood is “nonredundant.” The effective size of ego’s network may tell us something about ego’s total impact; efficiency tells us how much impact ego is getting for each unit invested in using ties. They conclude an actor can be effective without being efficient, and an actor can be efficient without being effective.

Again, in Hanneman & Riddle (2005), these authors describe that constraint is a summary measure that taps the extent to which ego’s connections are to others who are connected to one another. If ego’s potential trading partners all have one another as potential business partners, ego is
Related Content

**Older Adults with AMD as Co-Designers of an Assistive Mobile Application**
[www.igi-global.com/article/older-adults-with-amd-as-co-designers-of-an-assistive-mobile-application/107990?camid=4v1a](www.igi-global.com/article/older-adults-with-amd-as-co-designers-of-an-assistive-mobile-application/107990?camid=4v1a)

**Phraseology in English as an Academic Lingua Franca: A Corpus-Based Study of Prepositional Verbs in Writing by Chinese, American, and British Students**
[www.igi-global.com/chapter/phraseology-in-english-as-an-academic-lingua-franca/162967?camid=4v1a](www.igi-global.com/chapter/phraseology-in-english-as-an-academic-lingua-franca/162967?camid=4v1a)
Large Retailers’ Responsible Initiatives in Support of Local Communities
[www.igi-global.com/article/large-retailers-responsible-initiatives-support/71054?camid=4v1a](www.igi-global.com/article/large-retailers-responsible-initiatives-support/71054?camid=4v1a)

Student Perceptions and Adoption of University Smart Card Systems
[www.igi-global.com/chapter/student-perceptions-adoptions-university-smart/68278?camid=4v1a](www.igi-global.com/chapter/student-perceptions-adoptions-university-smart/68278?camid=4v1a)