Importance of First Steps in a Community for Growth, Diversity, and Engagement

Atef Chaudhury*
University of Waterloo
Waterloo, ON N2L 3G1, Canada
a29chaud@uwaterloo.ca

Myunghwan Kim
LinkedIn Corporation
Mountain View, CA 94043, USA
mukim@linkedin.com

Mitul Tiwari
LinkedIn Corporation
Mountain View, CA 94043, USA
mitultiwari@gmail.com

ABSTRACT

In both the online and o ine world, networks that people form within certain communities are critical for their engagement and growth in those communities. In this work, we analyze the growth of ego-networks on LinkedIn for new employees of companies, and study how the pattern of network formation in the new company a ects one's growth and engagement in that company. We observe that the initial state of ego-network growth in a newly joined company shows strong correlations with the future status in the company { such as network size, network diversity, and retention. We also present some key patterns that demonstrate the importance of the rst few connections in the new company as well as how they lead to the phenomena we observed.

1. INTRODUCTION

Social networks in a community like a company play a important role in individual's growth and engagement in that community { such as an individual's performance [2, 5, 6], personal regard [3], and the employee turnover [4]. In particular, when one forms a new social network after joining a new community, we can imagine that the growth of the ego-network within the new community (connections of an individual and the network among the connections) are crucial for faster growth and better engagement in this community.

Through the analysis of LinkedIn's company network data, we discover that the initial patterns of a new employee's ego-network have strong relationships with the future status of that employee in the company. Speci cally, network size, network diversity, and seniority of the rst few connections that a new employee makes are correlated with the network size, network diversity, and company retention rate of this new employee in the future. For example, the new employee initially connected with more senior employees is less likely to leave the company after a certain time.

To explain our discovery, we present here some key patterns in the ego-network growth regarding triangle-closing { well-known as a typical way of network formation. We demonstrate the propagation of triangle-closing, describing that one triangle-closing can open up the opportunity to close another triangle. For instance, an employee can meet another team's manager through his or her own manager, and later can be connected with this other team's members through their manager. By studying the triangle-closing propagation, we nd the importance of the rst few connections in the expansion of a new employee's network in the company. In particular, by combining the homophily on functional groups ($e.g.$ HR, engineering, etc.) and triangle-closing, we show the way of diversifying functional groups in one's ego-network through triangle-closing. These two mechanisms of the ego-network growth are aligned with our observations, described above.

*This work was done when he worked for LinkedIn.

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2889416.

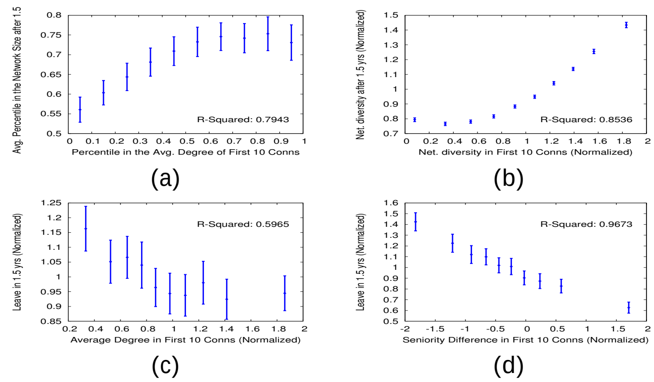

Figure 1: Relationships between a new employee's first 10 connections and the status of the employee after 1.5 years

2. OBSERVATIONS

Setup. We select top 500 companies in LinkedIn by their average degrees, and investigate pro le and the ego-network growth pattern of each employee in these companies. The number of employees in the top 500 companies is still over a million and even the new employees in 2013 is over 100k.

Analysis. From the top 500 companies, we study the employees who newly joined the companies in 2013 and have at least 20 connections within the companies. We then measure the following metrics from each of the rst 10 connections of the chosen new employees:

- Network size: the number of nodes in the ego-network
- Network diversity: the probability that a pair of nodes in the ego-network are in di erent functional groups
- Seniority di erence: the di erence of seniority level between a given employee and his/her connection

We use the standardized functional groups and seniority levels in LinkedIn, but omit the details due to the lack of space. For each new employee, we obtain the average value of the above metrics for the rst 10 connections, and normalize it by using percentile or dividing by the average value in the company to see the status within the company. We then compare the metrics from the rst 10 connections with the network size, network diversity, and retention rate after 1.5 years since each new employee joined the company.

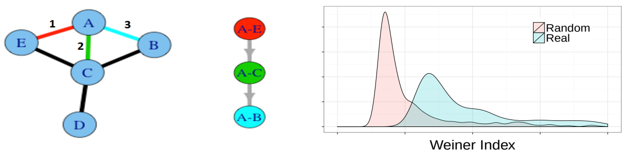

Figure 2: Triangle-closing propagation in the real-world

Figure 1(a) and (b) depict that the network size and diversity in the rst 10 connections' ego-networks is strongly correlated with the network size and diversity of the new employee's after 1.5 years. That is, as the rst connections have large and diverse ego-networks, the corresponding new employee is likely to have a large and diverse ego-network.

The rst 10 connections are also correlated with network-irrelevant status of a new employee after 1.5 years { retention rate. In Figure 1(c) and (d), for each new employee we compare the network size and seniority di erence of the rst 10 connections with the probability that the given employee leaves the company in 1.5 years. As a new employee is connected with more senior and well-connected (*i.e.* having a large ego-network) employees in the beginning, the new employee is less likely to leave the company in 1.5 years.

3. EGO-NETWORK GROWTH PATTERNS

While triangle-closing is regarded as one of the well-known patterns, it cannot solely explain the phenomena in the previous section. To further understand our observations, now we study ner patterns of ego-network growth.

Triangle-closing propagation. First we present the triangle-closing propagation. In Figure 2, suppose that A-E, A-C, and A-B are formed in turn on top of an existing network in black (left). Then, A-C closes a triangle A-E-C as well as creates another opportunity of triangle-closing A-C-B. Once A-B closes this triangle A-C-B, we view this as the propagation of triangle-closing. We de ne a E-Graph by representing a connection as a node and linking from the second connection of each triangle to its third connection (middle).

To verify the presence of the triangle-closing propagation, we represent an entire company network by E-Graph and compute the Wiener Index [1], which implies the average depth in a graph, for each connected component of the E-Graph. A high Weiner Index indicates chain structure, while a low Weiner Index means broadcasting structure. For comparison, we randomize all the timestamps of connections in the same company network and compute the Wiener Index in the same way. The right plot in Figure 2 shows the distribution of Wiener Index for each E-Graph of the both networks. The high trend for the real-world network as opposed to the randomized network indicates the existence of propagation behavior in the real-world.

We also investigate how much triangle-closing is contributed by each connection in the ego-network. We compute the effective volume of triangle-closing led by a given connection as follows. In the E-Graph representation, each leaf node (*i.e.* a connection in the original network) gives the credit to its parent node. However, if there exist multiple parent nodes (closing multiple triangles), we divide the credit equally and give the even value to each parent node. By running this process up to the root node, we aggregate such credits from the other E-Graph nodes. In the same way, we compute the e ective volume of triangle-closing with respect to each connection in the timestamp-randomized network.

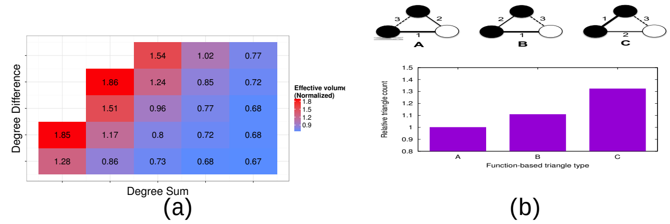

Figure 3: Ego-network growth pattern through triangle-closing

Figure 3(a) depicts the relative e ective volume of triangle-closing in the real-world networks compared to the random network. We group each connection by the degree sum and di erence between its two end users, and represent the relative value by each group. Note that high degree di erence and low degree sum (the connection between high-degree and low-degree users) has the most e ective volume of triangle-closing. As a new employee initially has a low degree, we see that a new employee's connections with the high-degree employees can lead to more triangle-closing, which will accelerate the growth of the new employee's ego-network. This agrees with our observation in Figure 1(a).

Group-based triangle-closing. Next we bring the functional group into the triangle-closing. In particular, we count the three kinds of triangle-closing in Figure 3(b). These types are identical in static networks but di erent if the order is considered. Type C, which expands the ego-network to another functional group via someone in the same functional group, is the most likely. This is aligned with our observation in Figure 1(b) in the following sense. The rst 10 connections can act like agents for the new employee's ego-network expansion to other functional groups, so the network diversity of the rst 10 connections can be correlated with the network diversity of the given new employee.

4. DISCUSSION AND FUTURE WORK

We presented relationship between a new employee's rst connections and the future status, and proposed some common patterns in ego-network growth that can lead to the phenomena we observed. Given that new employees initially tend to be connected with direct managers and team members, we conjecture that their status in company networks can in uence the new employees' growth in the company.

A natural next step then includes collecting more evidence about our conjecture and coming up with a model embracing all these common patterns as well as being able to explain our observations. We leave this for future work.

5. REFERENCES

[1] A. Anderson, D. Huttenlocker, J. Kleinberg, J. Leskovec, and M. Tiwari. Global diffusion via cascading invitations: Structure, growth, and homophily. In *WWW*, 2015.

[2] R. S. Burt. *Structural holes: The social structure of competition*. Harvard University Press, 2009.

[3] W. C. Carter and S. L. Feld. Principles relating social regard to size and density of personal networks, with applications to stigma. *Social Networks*, 26(4):323–329, 2004.

[4] T. H. Feeley, J. Hwang, and G. A. Barnett. Predicting employee turnover from friendship networks. *Journal of Applied Communication Research*, 36(1):56–73, 2008.

[5] R. Reagans and E. W. Zuckerman. Networks, diversity, and productivity: The social capital of corporate r&d teams. *Organization Science*, 12(4):502–517, 2001.

[6] T. Wolf, A. Schroter, D. Damian, and T. Nguyen. Predicting build failures using social network analysis on developer communication. In *ICSE*, 2009.