**Multilayer Network Analysis of Investment Patterns**

Jessica Santana (jsant), Raine Hoover (raine), and Meera Vengadasubbu (meerav)

December 11, 2014

## *SECTION 1. INTRODUCTION/MOTIVATION/PROBLEM DEFINITION*

Sociologists recognize that actors are likely to prefer familiar exchange partners, with which they have already exchanged previously, to unfamiliar exchange partners (Podolny 2001, Podolny 1994, Sorenson and Stuart 2001, Powell et al. 2005, Sorenson and Stuart 2007, Granovetter 1985) [1, 2, 3, 4, 5, 6]. Familiarity from prior exchange increases the actor's trust that the exchange partner will perform as expected, thus reducing the actor's uncertainty in the outcome of the exchange. However, analyses of repeated exchange, particularly in the context of startup investment, have been limited to inter-organizational exchange. Such analyses evaluate a firm's decision to invest in a company following prior investment. Some studies appeal to individual-level analysis of the investor (Sorenson and Stuart 2007) [5], but halt this precision in the case of the startup. There is much to learn about the structural relationship between investment decisions and the individual entrepreneur. Our research drills beyond the observation of repeat investment in a company, to analyze the longitudinal, dynamic possibility that investment is based on the individual entrepreneur *rather than* the company. We theorize that investment follows the entrepreneur, demonstrating a spillover effect from an entrepreneur's prior success. Confirming this theory would indicate that investors attribute a startup's success to the entrepreneur rather than the organization or the product.

Our hypothesis is that a startup that hires an employee from a separate startup that has received investment will also receive investment from the same investor. This behavior can be designated as an investor "following" an employee, or, more directly, an employee "being followed" by an investor. We use similarity metrics on a multiplex network to test our hypothesis. Multiplexity has previously been evaluated via proxies such as cohesion, or multiple distinct pathways between nodes A and B (Powell et al. 2005) [4]. Our research leverages innovative multiplex network analysis and network alignment techniques to retain the longitudinal, multidimensional complexity of investment-employment networks and directly compare the similarity of sub-networks across the layers of this multiplex network.

## *SECTION 2. PRIOR WORK*

Our analysis builds on three critical studies of multiplex network analysis. First, Kivela et al. 2014 [7] define multiplex networks as occurring when one network (layer A) is connected to another network (layer B) via a uniting set of nodes in common. We use this definition to generate our network, described below. Second, Szell et al. 2010 [8] demonstrates that distinct roles in a network can be analyzed as a multiplex network, where each role comprises a different layer linked together by the individual performing the given roles. The authors use online gaming community data to observe the relationship between a player's role in one layer of the network and their behavior in a separate layer of the network. Given these findings, we begin our analysis with the Jaccard coefficient, which can be used to measure similarity of relationships between two or more network layers. Third, we enhance our analysis beyond a simple Jaccard measure based on the findings of Meng et al. 2014 [9], which incorporates temporal dynamics in multiplex network analysis. The authors use frequency of communication types, namely digital and in-person communication, across timestamps of lengths varying from one week to six months to determine how similar various types of communication are across time. From this analysis, we learn that common edge, adjusted common edge (including the Jaccard index), Pearson correlation (of degree distribution), and graphlet degree distribution agreement can all be used to measure similarity of network layers. In addition, these authors find that period length and edge frequency thresholds influence network density. Expanding on Meng et al. 2014, we develop a novel coefficient that divides the Jaccard measure into each distinct network layer for comparison.

*SECTION 3. MODEL/ALGORITHM/METHOD*

**Crunchbase Dataset**

We collected data to construct our network from Crunchbase [10], an online directory of technology startups. This directory provides timestamped information on startup employment and investment, and includes hyperlinked profiles of startups, employees, and investors. Crunchbase is limited primarily to technology startups that have already received funding and are generally located in Silicon Valley. Profiles are generated and maintained by the public, in wiki fashion. This implies that Crunchbase data is prone to data entry errors, inaccuracies, and selected content inclusion. It is important to thoroughly clean this dataset prior to analysis. For example, we realized that Crunchbase often miscategorizes investors as employees. In these cases, investors are listed as employees of the startups in which they invest, either as board members or as investor team members. To address this, in our pairwise comparisons described below, we ignore any investor/employee pair that are in fact the same person.[1] Despite these limitations, Crunchbase is the primary public source of information on startup histories and networks. Having a profile in Crunchbase is a strong signal of legitimacy for entrepreneurs. Entrepreneurs and others listed in Crunchbase are motivated to ensure that their profiles are accurate. Moreover, Crunchbase caters to network researchers and is continuously improving the availability of historical network data.

Making use of CrunchBase's monthly export on investment and their public API, we created two edge lists describing 1) investment and 2) employment relations. We generate these networks through the lens of investment first, followed by employment. In other words, we first collect names of companies that received any investment along with the name of the investor and the date of investment. We then collect names of people who were ever listed as employees of those companies and the dates of their employment.

This approach restricts our analysis to those companies that have received investment, those employees that have worked for companies that have received investment, and those investments that are listed in Crunchbase. If any investor invested in a company that is not listed in Crunchbase, this investment is not included in our analysis. If an employee worked at any time for a company that did not receive a Crunchbase-listed investment, this employment is not included in our analysis. Thus, an employee that our algorithm registers as consistently being followed by an investor could in fact have worked in between the dates of employment seen by our algorithm for companies that never received investment. We do not, however, expect this issue to significantly affect our results. Crunchbase does not generally add a pre-IPO startup to its directory until that startup receives funding. The vast majority of those startups listed in Crunchbase, therefore, have received at least some funding.

**Multiplex Network of Investment and Employment**

Startup finance networks are complex, dynamic systems with multiple layers of nodes and edges. We use multilayer network analysis drawn from Kivel et. al to analyze the relationship between two startup finance network layers: investment and employment [7]. These networks share company nodes in common. In other words, investors invest in companies that employ employees. The graphs are aggregated over time; if an investor has ever invested in a company, there is an edge between that investor and that company; if an employee has ever worked at a company, there is an edge between that employee and that company. The edges of employment have the start date of employment as an attribute, and the edges of investment have the date of the investment as an attribute. If there are multiple dates of employment or investment, that edge's attribute is a list of dates. These networks are then modeled as a bipartite directed graph. In the investment graph, all edges point from investor to company. In the employment graph, edges point from employee to company.

[1] There are 242 such pairs.

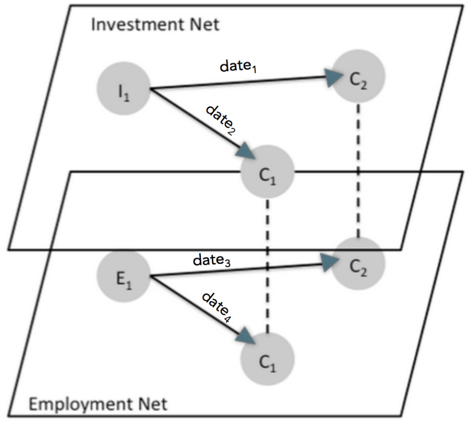

Figure 1: Company-Investor and Company-Employee multilayer network

We thus have two overlaid graphs with company nodes in common: $G_{I,C}$ (I, C) and $G_{E,C}$(E, C), where I is the set of investor nodes, E is the set of employee nodes, and C is the same set of company nodes.

**Exploratory Network Statistics**

The multilayer investment-employment network is composed of 69,598 total nodes. As of September 2014, the 14,786 companies in this network have employed 39,587 employees and have been invested in by 15,198 investors. 1,662 employees are also listed as investors.

| Table A: Entire Network | Table B: Company-Employee Directed Graph | Table C: Company-Investor Directed Graph |
|---|---|---|
| Companies: 14,786<br>Employees: 39,587<br>Investors: 15,198<br>Edges: 101,593<br>Nodes: 69,598 | Edges: 47,645<br>Nodes: 54,373<br>Employees: 39,587<br>Companies: 14,786 | Edges: 53,948<br>Nodes: 30,010<br>Investors: 15,198<br>Companies: 14,584 |

Table 1: Network statistics

**Initial Network Restrictions Made for Analysis**

The behavior of following and being followed would only be evident if there was a prior history of employment or investment. Therefore, in our calculations we did not take into account employee or investor nodes that have less than two edges. Of the 15,198 investors in our network, 4,092 have invested more than once. Of the 39,587 employees in our network, 6,647 employees have been employed by more than one company.

**Method of Analysis: Network Similarity**

We estimate that the structure of an investor's network will ultimately resemble that of an employee's network as the investor invests in each of the employee's subsequent employers. This can be understood as a lagged mirror effect, in which the investor's network mirrors the employee's network with a lag demonstrating the order in which investment follows employment. Using our graph model, we can analyze the presence or absence of this mirror effect by investigating the shared edges between an investor node and an employee node-- companies in which the investor has invested and where the employee has worked--and the date attributes on those edges. If the mirror effect is present, the number of shared edges, where the date of employment precedes the date of investment, will be non-zero.

In other words, if our hypothesis holds, the investor's network will be similar to the employee's network, and that similarity will take into account the dates on the common edges. We measure pairwise similarity using two metrics: 1) Jaccard Coefficients and 2) what we will call intersect proportions.

Both of these methods begin by analyzing the intersection of the investor and employee networks. Given an investor Ii and an employee Ej, the intersection encompasses the edges to the company nodes that are shared between Ii and Ej. These are the companies that both employ Ej and receive funding from Ii. We impose a further requirement to qualify an edge as shared, in order to capture the temporal element of our inquiry: an edge is only counted in the intersect of Ii and Ej if employment occurred prior to investment. If there are multiple dates of employment or investment, we use the earliest date. To ensure that this intersection represents a "follow", the size of this set of intersects should be greater than 1, implying there was more than one time that this investor invested in a company at which Ej worked. We name the cardinality of the set of intersections thus defined as p.

Jaccard Coefficient

The Jaccard index is a measurement of pairwise similarities between vertices, and is calculated by taking the intersection of two vertices' edge sets divided by the union of the two vertices' edge sets [11]:

$$\sigma_{\text{Jaccard}} = \frac{|\Gamma_i \cap \Gamma_j|}{|\Gamma_i \cup \Gamma_j|}$$

Meng et. al [9] use this index to measure the similarity of the different media layers of a communication network among college students. The layers in Meng et. al are node-aligned; the same college student wedded to the same node ID is present in every layer, and all nodes are of the same category, students talking to other students. In our graph, the nodes in the two layers are heterogenous; we have investors and employers, two distinct categories, and though they are both connected to company nodes that have the same ID in the two layers, these connections are of very different natures. Specifically, investors have higher degrees than employees; investors invest in more companies than employees are employed. In the Jaccard calculations, therefore, the denominator (the union of edges in either the employee or the investor) is quite large in comparison to the numerator (the number of edges in both), and Jaccard indices are quite low.

Intersect Proportions: p/m and p/k

In order to find a measure with more spread, with the aid of Himabindu Lakkaraju we made a slight modification to the Jaccard method that gave us more meaningful data. We kept the numerator as p, the edges in the intersect that abided by the heuristic that employment preceded investment, but made the denominator relative to the type of node we were investigating. For investor nodes, the denominator is k, the total number of investments made by that investor, and for employee nodes, the denominator is m, the total number of employments held by that employee.

p/k indicates the similarity of the individual investor and employee nodes. A high value of p/k indicates that, of the companies investor Ii invested in, a majority of them were companies in which employee Ej had worked prior to Ii's investment and can be expressed as following

$$p/k\,(Ii,\, Ej) \;=\; \frac{|O(Ii) \cap O(Ej)|}{|O(Ii)|} \quad \textit{where } O(X) \textit{ is set of out nodes of } X$$

p/m indicates the similarity of the individual employee and investor networks. A high value of p/m indicates that of the companies employee Ej worked for, a majority of them were companies that Ii had funded after Ej's employment and can be expressed as following

$$p/m\,(Ej,\,Ij) \;=\; \frac{\left|\begin{matrix} O(Ii) \\ \bigcap \\ O(Ej) \end{matrix}\right|}{|O(Ej)|} \quad \textit{where } O(X) \textit{ is set of out nodes of } X$$

Algorithm

We now describe the process of translating these pairwise comparisons into meaningful metrics for the graph as a whole.

For each investor, we calculate a set of similarity metrics for each employee as Mi. For each employee, we calculate a set of similarity metrics for each investor.and can be expressed as Mj. Both of these can be expressed as follows:

$$Mi \;=\; \{p/k\,(Ii,\ Ej)\,|\,p/k \;>\; 0.0\}$$
$$Mj \;=\; \{p/m\,(Ej,\ Ii)\,|\,p/m \;> 0.0\}$$

For each pairwise similarity metric, each employee Ej, then, has a value for that metric with each investor Ii stored in Mj, and each investor Ii, then, has a value for that metric with each employee Ej stored in Mi.

Then investors following at least one employee at any threshold can be defined as:

$$I_{1Ft} = \;\{Ii \mid\; any\ of\ p/k(Ii) \;>=\; t\,\}$$

And employees followed by at least one investor at any threshold can be defined as:

$$E_{1Ft} = \;\{Ej \mid\; any\ of\ p/m(Ej) \;>=\; t\,\}$$

We then tally up the number of matches in Mi for whom the value of the similarity metric is above a certain threshold, where the threshold varies between 0 and 1 (the minimum and maximum values for the similarity metrics chosen). With the similarity metric p/k, q is the number of investors for whom this tally is non-zero, and indicates the prevalence of "following," i.e. where an investor funds a company after a previously funded employee joins. With the Jaccard similarity metric, we call this number qj.

We repeat this process for each Mj, and call this number s with the similarity metric p/m, a metric that indicates the prevalence of “being followed”, i.e. where an employee changes employers, and an investor that invested in a prior employer invests in their new employer. With the jaccard similarity metric, we call this number sj.

The final metrics calculated are: 1) qI, the proportion of investors for whom this “following” behavior is present at the level of the given threshold, q/|I|, where I is the set of investors with more than one investment, 2) qI’, the proportion of investors who have some “following” behavior for whom this behavior is present at the level of the given threshold, q/|I’| where I’ is the set of investors with an employee match with whom they have a non-zero similarity metric, 3) sE, the proportion of employees for whom this “being followed” behavior is present at and above the level of the given threshold, s/|E|, where E is the set of employees with more than one edge, and 4) sE’, the proportion of employees who have some “being followed” behavior for whom this behavior is present at and above the level of the given threshold, s/|E’|, where E’ is the set of employees with an investor match with whom they have a non-zero similarity metric, 5) sj, the proportion of investors out of all investors who have some “following” behavior for whom a match is present with a Jaccard value at or above the given threshold, 6) qj, the proportion of employees out of all employees who have some “being followed” behavior for whom a match is present with a Jaccard value at or above the given threshold.

This can be defined as

$$qI \;=\frac{|I_{1Ft}|}{|I|} \quad sE \;=\frac{|E_{1Ft}|}{|E|} \quad qI’ \;=\frac{|I_{1Ft}|}{|I’|} \quad sE’ \;=\frac{|E_{1Ft}|}{|E’|} \quad qj \;=\frac{|I_{1Ft}|}{|I’|} \quad sj \;=\frac{|E_{1Ft}|}{|E’|}$$

For example, for a specific employee Ej, if an investor Ii has invested in 10 companies, 5 for which this employee worked (and was working before the time of Ii's investment), this match would contribute to the tally for Ej's Mj, and Ej would be included in the tally q for all employees with investor matches at a threshold value of 0.5 or less.

```
#-- calculate similarity metrics for each employee
initialize pMmap, key → employee node ID, val →  list of p/m's for each investor
for every employee node e:
   for every investor node i:
         intersect = 0
                  get the neighbors of e and i, get the edges in their intersect
                  for each edge in the intersect:
                           if dateEmployment < dateInvestment
                                    intersect += 1
         pm = intersect/(total # of e's edges)
         if e in pMmap:
                  pMmap[e] = pMmap.append(pm)
         else:
                  pmMap[e] = [pm] # make pm first element in list

#-- tally number of employees (investors) with at least one match at or above threshold
nonzeroMatches = {}
for key in pmMap:
         count = 0
         for metric in pmMap[key]:
                  if metric >= threshold:
                           count += 1
         if count > 0:
                  nonzeroMatches[key] = count
se = (size of nonzeroMatches key set at this threshold / # of employees with degree > 1)
se' = (size of nonzeroMatches key set at this threshold / # of employees with
         one nonzero pm)
```

Figure 2: Pseudocode for s, without loss of generality for q, Jaccard

### Algorithm refinement

We began with pairwise comparison, as described above, of every investor that had more than one investment with every employee who had more than one employment, for a total number of 29,359,799 comparisons. In an effort to improve the efficiency of our algorithm, we took a new approach that utilized the aforementioned single layer graph with dated edges from companies to employees and investors. Rather than compare all investors with all employees, we iterated over the companies themselves to find pairs of investors and employees who, by this mode of discovery, were guaranteed to have at least one shared edge. With this tweak, we successfully decreased the number of comparisons by several orders of magnitude to only 18,073. Thus for analysis of pairwise similarities in multilayer networks, collapsing the layers and iterating over the nodes that are present in both sets improves the efficiency of the algorithm considerably.

```
#-- efficient/improved algorithm
for every company:
         get employee and investor neighbors
         for each e in employee neighbors:
                  for each i in investor neighbors:
                           if we haven't already seen e,i:
                                    carry out pm calculations as in first algorithm
```

Figure 3: Pseudocode for refined algorithm

## *SECTION 4. RESULTS AND FINDINGS*

### Results

The results for different thresholds are summarized in Table 2 and Figure 4. The proportions over the entire set of employees or investors, q/|I| and s/|E| respectively, are low, falling in the range of .052 at the lowest threshold and .007 at the highest threshold for q, and .105 at the lowest threshold and .057 at the highest threshold for s.

However, once we limited our denominator for the proportion to just those employees and investors who had some nonzero amount of follow/following behavior, the numbers increased dramatically: q/|I'| is .566 at the lowest threshold, and .074 at the highest threshold, while s/|E'| is .998 at the lowest threshold and .540 at the highest threshold. These s/|E'| values state that in the set of employees that are at all followed by investors, 99.8% of these employees are followed 10% of the time, and 54.0% are followed at the highest value of p/m, 1.0--at every company they have worked, this investor invested after they joined that company. Jaccard statistics are low and follow the same shape.

| threshold | qI' | sE' | qj | sj | qI | sE |
|---|---|---|---|---|---|---|
| 0.100000 | 0.566449 | 0.998382 | 0.514161 | 0.396440 | 0.052052 | 0.104861 |
| 0.200000 | 0.331155 | 0.996764 | 0.263617 | 0.218447 | 0.030430 | 0.104691 |
| 0.300000 | 0.228758 | 0.982201 | 0.119826 | 0.100324 | 0.021021 | 0.103161 |
| 0.400000 | 0.189542 | 0.967638 | 0.091503 | 0.077670 | 0.017417 | 0.101632 |
| 0.500000 | 0.154684 | 0.938511 | 0.039216 | 0.043689 | 0.014214 | 0.098572 |
| 0.600000 | 0.126362 | 0.854369 | 0.000000 | 0.000000 | 0.011612 | 0.089735 |
| 0.700000 | 0.076253 | 0.593851 | 0.000000 | 0.000000 | 0.007007 | 0.062373 |
| 0.800000 | 0.076253 | 0.561489 | 0.000000 | 0.000000 | 0.007007 | 0.058973 |
| 0.900000 | 0.074074 | 0.540453 | 0.000000 | 0.000000 | 0.006807 | 0.056764 |
| 1.000000 | 0.074074 | 0.540453 | 0.000000 | 0.000000 | 0.006807 | 0.056764 |

Table 2: Threshold and corresponding graph level similarity metrics

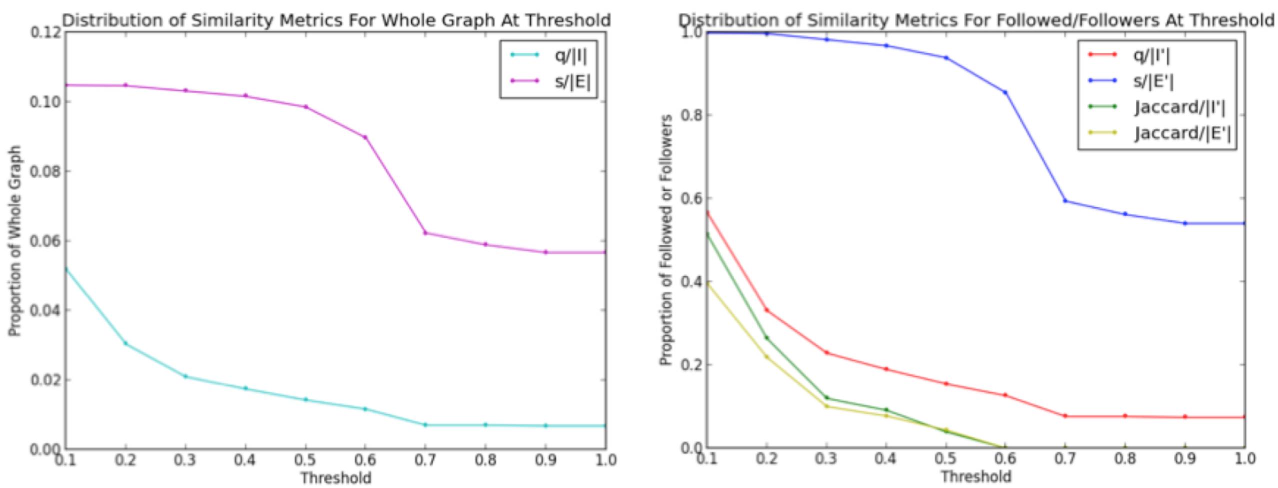


Figure 4: Distribution of graph level similarity metric results for different threshold values in proportion to whole graph (left) and graph that displays followed/following behavior (right)

The distribution of the pairwise intersect proportions can be seen in Figure 5. The intersect proportion relative to the investor, p/k, is very low in a majority of the pairwise comparisons made, while the intersect proportion relative to the employee, p/m is high (1.0) in a large portion of the pairwise comparisons made.

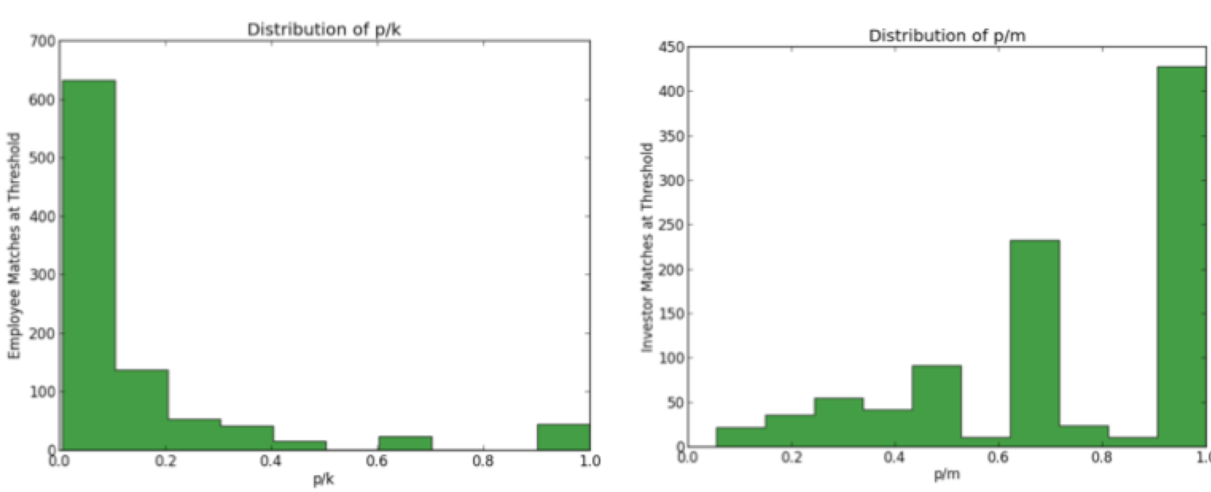


Figure 5: Distribution of pairwise intersect proportions.

Figure 6A depicts the number of employees with at least one match for p/m at the given threshold, where the threshold is incremented in steps of .01 in order to drill further down into the data. The graph is stepwise, with the most dramatic step occurring at .65. Below .65, there is much more variation in the number of employees at each threshold, but above .65, the graph stays close to 334-367. Furthermore, we can see from the graph that a small but significant number of employees, 334, are followed with an intersect proportion of 1.0. These patterns are maintained when we plot the employees with not just one investor match, but all investor matches at this threshold in Figure 6B. The number of employees who are followed by all their investors at a threshold of 1.0 is 326, very close to the number of employees with at least one investor match.

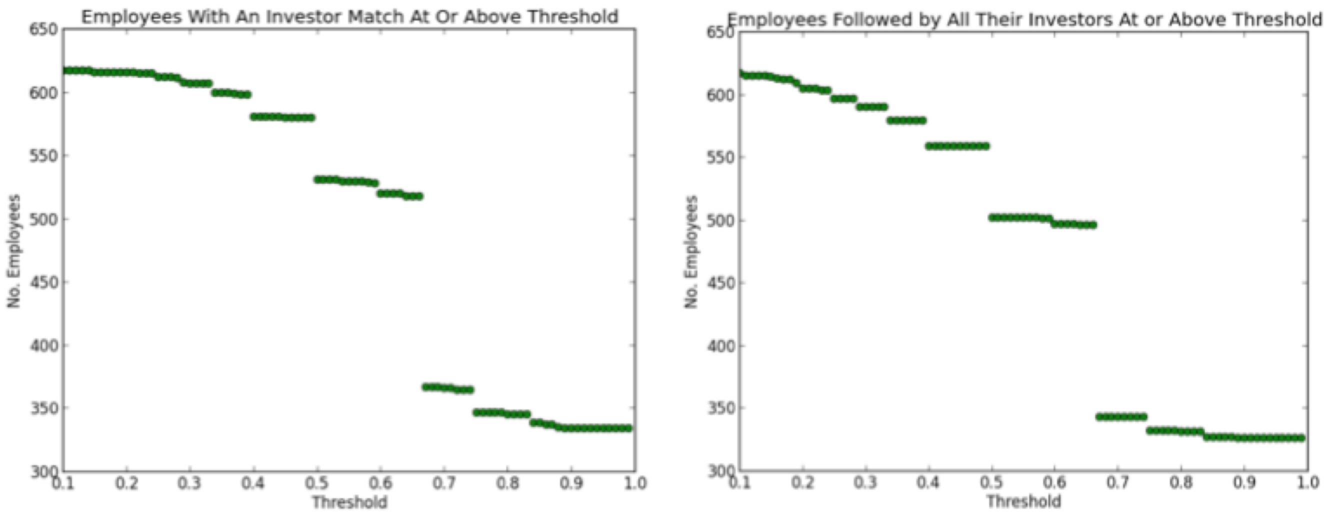


Figure 6A: Count of employees with at least one investor match at or above threshold, Figure 6B: Count of employees with all investor matches at or above threshold

To gain more insight into the 326 employees with a 1.0 similarity coefficient with all of their investors, we obtained the distribution of number of investors following that employee, displayed in Figure 7. We find that while a majority of this number is coming from employees with only one investor follow, there are employees for whom two, three, four, five, and even six investors follow them consistently, across all of their ventures, with a 1.0 p/m value.

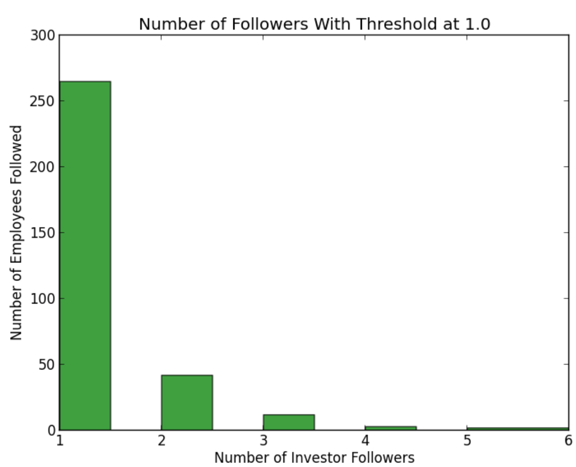


Figure 7: Number of investors consistently following employees at threshold 1.0

**Findings**

While the numbers for q/|I| are low compared to s/|E|, as with the numbers for q/|I'| and s/|E'|, this is to be expected. Investors may follow multiple employees, but our p/k values only investigate the intersection with a single employee; out of their total investments, those in common with one employee's employment would not be a large number. For each individual employee, however, the ratio of the intersect to the total number of that employee's employments would be high. This is reflected in the distributions of the intersect proportions: employees have a relatively high instance of high p/m values, while investors have a relatively low instance of high p/k values. Following this reasoning, we take s/|E| and s/|E'| to be more informative of the nature of the following/followed phenomenon in the graph. Even so, it is worth noting that 5.20% of all investors in the graph follow an employee match with a .10 similarity value for p/k.

The low proportions of s/|E| for every threshold illustrate that as a proportion of the entire graph, the number of employees who show some behavior of being followed is low. Yet the idea that not every investor follows every employee is not surprising. Some employees have many changes in employment not because they are serial entrepreneurs, but because they are serial bad employees, and we would not expect an investor to follow these employees. As for the investors, some are large companies themselves, that make investments in a variety of startups, and do not have personal or even very deep professional connections with the employees of the startups in which they invest. Even with these factors that would minimize the phenomenon in the unrestricted set of employees with more than one edge, the data reflects that even out of the entire set of employees, 10.163% of employees are followed by at least one investor at a p/m value of .40 or higher.

q/|I'| and s/|E'| represent the amount of investor following and employees being followed when the analysis is restricted to those investors that show some amount of following behavior, and those employees that show some amount of being followed. The numbers for s/|E'| are much higher. Indeed, more than half of these employees show an investor that follows them 100% of the time. Furthermore, the small spread observed in Figure 6 for thresholds above .65 implies that if an employee is followed above the threshold .65, they are likely followed at thresholds above .65. This result, coupled with a count of 334 employees that have at least one investor match with a p/m value at the highest threshold 1.0, suggests that there is a subset of the graph for which this being followed behavior is a significant phenomenon, as predicted by our hypothesis. s/|E| tells us that this subset of maximum similarity "follows" makes up 5.88% of the graph.

The 326 employees with 1.0 matches across all investors indicate that the behavior of being followed not by one but by all investors is prevalent. The distribution of the number of investors for each of these employees illustrates that while a majority of this subset only have one investor (and thus by default all investors following at or above the threshold), there is a small, elite group of employees in the graph who have multiple investors consistently following them with the highest similarity metric.

**Evaluation of methods**

The Jaccard coefficient, while not very informative above a certain threshold value, is a good sanity check for the intersect proportions. p/(k+m), the equivalent of the Jaccard coefficient, normalizes the size of the intersect in relation to the maximum possible overlap, the total network of both Ej and Ii. The whole graph similarity metric| values using the Jaccard coefficient were extremely close in value and had the same shape, confirming that p/m and p/k are getting at similar phenomena from different angles.

**Conclusion**

In this study, we used multiplex network analysis and network alignment techniques to evaluate the presence of investor "following" behavior, where an entrepreneur receives investment from the same investor as the entrepreneur moves from company to company. We hypothesized that a startup that hires an employee (i.e. entrepreneur) from a separate startup that has received investment will also receive investment from the same investor. We built a multiplex network of investment and employment relationships and applied pairwise Jaccard and "intersect proportion" similarity metrics to evaluate our hypothesis. We found that, after limiting our analysis to the set of employees that are at all followed by investors, 99.8% of these employees are followed in 10% of their employments, and 54.0% are followed in

100% of their employments. These results suggest that a subset of entrepreneurs are indeed followed by the same investors.

There are several potential extensions of this research. First, one could explore the patterns associated with those employees that are followed by investors 100% of the time. Perhaps there are interesting features of these employees that promote investor following, such as common affiliations with the investor or signals of high status. Secondly, the distribution of investor counts for consistent 1.0 investor "follows" in Figure 7 raises interesting avenues for research. The tail on the distribution suggests that the number of investors following at the maximum p/m value might fit the model of preferential attachment of investor followers to employees followed; it could be that the more investor followers at or above a given threshold an employee has, the more followers at or above that threshold an employee would gain. A third extension could be to further limit the denominators of our analysis, such as the investor set. Perhaps isolating the analysis to angel investors, for example, would reveal an increase in investor following, assuming that a more personal interaction between angel investors and entrepreneurs (in contrast to more hierarchical institutional investment) would promote loyalty between the angel investor and the entrepreneur.

## References

[1] Podolny, Joel. M. 2001. "Networks as the Pipes and Prisms of the Market." *American Journal of Sociology*, Vol. 107, No. 1 (July 2001), pp. 33-60.

[2] Podolny, Joel M. 1994. "Market Uncertainty and the Social Character of Economic Exchange." Administrative Science Quarterly 39:458–83.

[3] Sorenson, Olav and Toby E. Stuart. 2001. "Syndication Networks and the Spatial Distribution of Venture Capital Investments." The American Journal of Sociology, Vol. 106, No. 6 (May, 2001), pp. 1546-1588

[4] Powell, Walter W., Douglas R. White, Kenneth W. Koput, and Jason Owen-Smith. 2005. "Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences." The American Journal of Sociology, Volume 110 Number 4 (January 2005): 1132–1205

[5] Sorenson, Olav and Toby E. Stuart. 2007. "The Evolution of Venture Capital Investment Networks." Unpublished Manuscript.

[6] Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." American Journal of Sociology, Vol. 91, No. 3: 481-510.

[7] Kivela, Mikko, Alex Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. 2014. "Multilayer Networks." *Journal of Complex Networks.* http://comnet.oxfordjournals.org/content/2/3/203

[8] Szell, Michael, Renaud Lambiotte, and Stefan Thurner. 2010. "Multirelational organization of large-scale social networks in an online world." *arXiv.* http://snap.stanford.edu/class/cs224w-readings/szell10multirelational.pdf

[9] Meng, Lei and Milenković, Tijana and Striegel, Aaron. 2014. "Systematic Dynamic and Heterogeneous Analysis of Rich Social Network Data." *Springer International Publishing*. http://dx.doi.org/10.1007/978-3-319-05401-8_3

[10] www.crunchbase.com

[11] Leicht, E. A. and Holme, Petter and Newman, M. E. J. 2006 "Vertex similarity in networks", *Physical Review,* Department of Physics, University of Michigan, Ann Arbor, Michigan. http://journals.aps.org/pre/pdf/10.1103/PhysRevE.73.026120