VCRank: A Data-driven Ranking of Venture Capital Firms

Siddharth Reddy SID.REDDY@CORNELLVC.COM
Cornell Venture Capital

Abstract

...

1. Motivation

Our motivation for this project is three-fold:

- Produce a marketing tool that will put CVC "on the map"
- Provide useful recommendations and comparisons to New Atlantic Ventures
- Generate novel visualizations of venture capital investment activity

2. Problem Setup

Data: As part of this project we plan to use several sources of venture capital investment data: Crunchbase, AngelList, Mattermark, Seed DB, Midas, NVCA, and Find The Best.

Objects: Venture capital firms (composed of individual investors), start-up companies

Observed Links: Investment rounds, exits (including acquisitions)

3. Problem Statement

We aim to rank venture capital firms by the following general criteria:

- Ability to pick successful start-up companies to invest in
- Ability to help portfolio companies grow
- Ability to attract other VC firms as co-investors
- Ability to attract and retain successful individual investors

This work is done for a Cornell Venture Capital project.
NMI5000 Spring 2014

4. General Approach

We plan on using several approaches to rank VCs:

Investor Rank: Here, we model co-investments as a random walk on a graph with VC firms as vertices. We draw edges between VCs when they invest in the same start-up company. Edge probability is measured as $e_{ij} = \frac{\text{\# times VC i co-invests with VC j}}{\text{\# times VC co-invests with other VCs}}$. The stationary probability given by the Pagerank algorithm gives us our ranking. This approach was inspired by a TechCrunch article describing a similar method [1]. We take our inputs from the Crunchbase data set.

EarlyBird Rank: Here, we compare VC firms' ability to pick successful start-up companies by comparing *when* any two firms invest in a company. Presumably, the VC firm that invests earlier than another firm (in the same start-up company) is better able to predict companies' ability to raise funds (which we use as a proxy for "success"). We take our inputs from the Crunchbase data set.

Growth Rank: Here, we measure how well a VC firm helps start-up companies grow by calculating the following metric for all (VC, Startup) pairs: $M1 = \frac{\text{Total Funding Raised}}{\text{Initial Investment Amount}}$. We take our inputs from the Crunchbase data set.

People Rank: Here, we measure how well VC firms are able to attract and retain successful individual investors. First, we calculate the following metric for all individual investors:

$.Then, we calculate the following metric for all VC firms:$

. We take our inputs from the AngelList data set.

We plan on generating the following visualizations:

- Clustering of individual investors by `Markets Seeking` and `Locations Seeking` in the AngelList data set
- Heat map of VC activity (i.e. funds raised by start-up companies from VC firms) by geographic region and

[1] http://techcrunch.com/2011/05/25/top-10-vc-firms-investorrank/

industry over time