The Role of Social Media in Financing of Technology Ventures

Aleksandar Bradic
Center for Financial and Management Studies
University of London, SOAS
abradic@acm.org

Abstract

In this research, we address the problem of estimating the future financing of technology ventures using publicly available data. In particular, we focus on the problem of estimating the hazard rate of future investment and exit options as a function of feedback and sentiment towards venture or product visible in the social media. We perform the analysis across multiple markets within Information Technology sector and aim at determining 'sensitivity' of VC investments to social feedback in each individual market. In doing so, we rely only on publicly available data and hope to come up with the model that can be used for estimation of investment prospects for new ventures.

Introduction

Venture Capital represents the most common mean of obtaining financing for new technology ventures. It is estimated that, only in the first quarter of 2011, Venture Capitalists have invested $5.8 billion across 736 deals, out of which 37% in the areas of Software, IT Services, Media and Entertainment (*PricewaterhouseCoopers Money Tree Report, 2011)*. Venture Capitalists usually employ "stage financing" by investing in portfolio companies across multiple rounds, between which progress and potential of the venture is evaluated and decisions are made on the preferred exit routes. By doing so, Venture Capitalists "vote" on the potential of new venture and select the exit route that maximizes expected return on investment. Given decisions are primary based on company's internal metrics such as user adoption, rate of growth, cash flows and secondly based on external factors such as industry and market conditions.

The subject of Venture Capitalist's decision making regarding timing and structure of financing in light of complete information about company's prospects has been well researched in the literature. However, it might be of interest to attempt at estimating likelihood of obtaining the next round of financing for new venture assuming only publicly available information. This scenario would correspond to the case of external venture capitalists, competitors and public markets attempting at estimating prospects of success for the new venture. In this regard, the existing

body of research does not provide sufficient number of explanatory variables that would enable us to establish such a relationship.

In this research, we argue that in the case of Technology companies, such variables might exist in the form of Social Media data. Given the nature of the process in which users and enterprise customers adopt new technology products, it is expected that sufficient trace of this activity can be found in the Social Media and can be used as a convenient proxy for some of company's internal metrics such as growth and user adoption. In particular, we test whether feedback volume and user sentiment present in social media can be correlated with future financing of new ventures. Given that Venture Capitalists are likely to continue investments in the high-growth high-potential companies, we expect this relationship to hold, especially in the case of Internet Consumer start-ups.

The results of such findings would be very valuable both for entrepreneurs, as a mean of evaluating growth and expansion potential of competitors in the same market and Venture Capitalists as a mean of screening the markets for new investments. Finally, any insights on the impact of social media to future investments and exit outcomes in particular industries might provide a guideline for new companies on whether they should focus on monitoring and affecting social media coverage in order to increase likelihood of continued financing and preferred exit option.

Research Questions

The goal of this study is to attempt at establishing relation between timing and structure of Venture Capital financing and respective user feedback and sentiment present in the social media.

In particular, this study aims at answering the following research questions:

- Can statistically significant relationship be established between investments in technology startups and respective coverage in social media?
- What is the nature of this relationship across different market segments?

Literature Review

Venture Capital plays a critical role in innovation cycle by providing financing for early-stage, high-potential, high risk, growth startup companies. Such companies find obtaining financing through traditional mechanisms virtually impossible due to the four critical factors: uncertainty, asymmetric information, the nature of firm assets, and the conditions in the relevant financial and product markets (Gompers and Lerner, 2006). Venture Capitalists aim at addressing some of these issues by playing the role of informed screening agents, providing superior evaluations of

project quality, taking active role in company development and monitoring of company prospects and performance (Metrick and Yasuda, 2010). By employing staged financing on their investments, VCs are able to control the risk and mitigate associated moral hazard issues (Wang and Zhou, 2002). Investments are spread across multiple rounds between which investors evaluate the prospects of the venture. Each financing round presents VCs with an option to assess possible exit routes, in terms of IPO, trade sale or liquidation (Giot and Schwienbacher, 2005) and take appropriate actions, reflected in continued financing and size of investment rounds. Factors affecting structure, timing and size of financing rounds are primary company growth, age, investment volume, and industry conditions (Gompers 1995), but also include factors such as geographic distance between VC investor and the firm (Tian 2011), trust (Bottazzi et al. 2010), patent applications (Haeussler et al. 2009) and state of public markets (Gompers et al. 2007). However, to the best of our knowledge, no research had been conducted regarding impact of the consumer feedback and online media coverage to venture capital financing.

Until recently, performing such analysis at large scale using publicly available data was virtually impossible. However, the advent of the Social Media has dramatically changed the ways in which users discover and interact with new products. It is estimated that over 75% of Internet users use "Social Media" by joining social networks, reading blogs or contributing reviews (Kaplan and Haenlein, 2010). By doing so, users freely express their sentiment in the way that leaves publicly available trace that can be used for such analysis. A number of publications have already established power of social media in predicting real-world outcomes such as federal elections (Tumasjan et al. 2010), box-office revenues for movies (Asur and Huberman 2010) and changes in market indexes (Bollen et al. 2010). To the best of our knowledge, no similar research has been conducted in the area of using social media to predict development and success of new ventures.

In support of this hypothesis, recent research on the usage of microblogging (Jensen et al. 2009) shows that 19% of the microblogs contain mention of the brand, out which 20% containing some expression of sentiment, with 50% being positive and 33% being critical of the company or product. This suggests that it should be feasible to extract genuine user sentiment towards new products from given sources. Additionally, social networks are more and more used by professionals (Skeels and Grudin, 2009) and as business information tools (Weiss, 2010). Also, a number of Venture Capitalists are starting to actively use social media to directly express their sentiment towards markets and new ventures (Teten and Farmer, 2010).

Finally, in the recent years, social media seems to have changed the patterns of consumption (Webster 2010) and represents one of the primary means of new product discovery, especially in the area of technology products. This suggests that a significant data should exist in social media that could be related to the outcomes of new technology ventures.

Procedure and Methods

In order to conduct proposed research, we aim at creating a unique dataset composed of existing data sources and publicly available information. Data required for this research can be classified in two categories: venture financing information and social media data.

As main source of technology companies, we use data from the Crunchbase portal (http://www.crunchbase.com/), representing open database of technology companies. This database provides detailed overview of 71,473 startup companies, 6,881 financial organizations and 23,229 funding rounds. A number of research projects such as (Block and Sandner, 2009) and (Dolencic, 2010) have already used this dataset to address various issues regarding venture capital investing. However, even though this database provides good source of technology companies, given the "open" nature of it's content, it is possible that some of the investment round information it might be incomplete and even incorrect. Recent research regarding consistency and reliability of Venture Capital Databases (Maats et al. 2011) indicates that significant discrepancies exist even between commercial VC investment databases such as VentureXpert (owned by Thomson Reuters) and VentureSource (owned by Dow Jones). Therefore, it seems that in order to obtain the most accurate information on financing for a particular technology company, all of the given sources must be consulted. Taking into consideration that access to VentureXpert and VentureSource databases is not available to individuals, one of the challenges of this research would be finding ways of obtaining access to this data.

Given the volume of data that has to be collected, potential number of sources that have to be consulted and limited time for conducting research, it seems that taking each technology startup into account as a part of the study is not feasible. Therefore, we decide to focus on the analysis of the sample of companies, using proportional random stratified sampling as data selection methodology.

The sample used in the research will be obtained by random sampling of companies from Crunchbase database, by selecting a proportional sample of companies within each technology category (Advertising, Consumer Web, eCommerce, Entertainment, Search, etc.). In order to ensure that enough public information exists for the companies in the sample, we restrict sampling to the set of companies that have already received Series A Venture Capital funding. For each company in the sample, we consult Crunchbase, VentureXpert and VentureSource databases and compose information about all financing rounds that given company has invested to date. In the case that inconsistencies between databases arise, we will consult external data sources (such as S-1 filling data in the case of companies that have went through the IPO process).

In sample selection, we only take in account companies that have received Series A funding during the historical time window for which sufficient amount of social data

exists. Determining the exact interval would be subject of the research, but we expect this window to range between last 4 to 10 years. Given that most Venture Capitalists time the exits of portfolio companies within 5 years period, described sample should be sufficient to cover all stages of venture financing.

For each of companies in the sample, we collect social media data across two categories: online news coverage and user feedback.

In order to generate online news coverage data, we collect number of times each company or related product appears in online media between two investment rounds. We split obtained counts between two figures: number of posts with positive and negative sentiment. In order to ensure that only authoritative media sources are taken into account, we use Google PageRank (Page et al. 1999), representing measure of "importance" of given source on the Internet and select only news from sources with PageRank value above certain threshold.

In order to generate user feedback data, we use a number of "social media" sources, across several categories such as: Blogs (Blogger, WordPress), Microblogging (Twitter, Indenti.ca), Social News (Digg, Redit) and Social Bookmarking (Delicious, StumbleUpon). For each of given categories, we generate counts and sentiment of posts regarding given company or product between each two financing rounds. One of the challenges in this process would be filtering of "spam" content from the data, and specific data cleaning methods would have to be used for each source.

For assisting the data collection process, a simple software tool will be created that will deal with preliminary data selection, filtering and sentiment detection. All of the results of this process will be manually reviewed in order to ensure the quality of the obtained data.

Finally, as a result of given process, a dataset will be created containing complete information about investment rounds for selected set of technology ventures along with total counts of positive and negative feedback from different types of social media between each two investment rounds.

Data Analysis

We perform quantitative study based on the obtained dataset in order to attempt to establish relationship between VC financing and social media sentiment. In this process, we use regression analysis with VC financing representing endogenous and social media representing exogenous factors.

We define funding amount and duration of given venture financing round as dependent variables. In order to deal with the fact that investment round data is right censored, we use methodology described in (Gompers, 1995) in estimating appropriate investment duration model.

Explanatory variables in given model include: technology category given company belongs to, count of positive and negative references regarding given company in online news, positive and negative sentiment count across different social media categories, company age and previous investments. Additionally, in order to offset for market effects, we use Dow Jones Venture Capital Index as explanatory variable in the model.

As a result of given estimation, we will be able to address the regarding validity of overall model and significance of factors associated with individual explanatory variables. After validating all of the assumptions of used regression methodology have been satisfied, we will be able to draw appropriate conclusions and interpretations of the impact of individual variables to the duration and size of investments.

Timescale, Resources and Budget

Given research should be conducted over the course of seven to eight months. Timeline of the projects is split across several phases:

1. Identification and analysis of data sources (Months 1 - 2)

During this period, an extensive survey of available data sources should be conducted, especially in the area of social media and selection of most relevant sources should be finalized. Initial analysis of data from selected sources should be conducted and potential data quality issues should be addressed.

2. Data Collection and Literature Review (Months 3 – 4)

During this period, described software tool for data aggregation should be developed and data collection should begin. All of the results should be manually reviewed before committing to the main body of data used for analysis. Throughout this period, intensive survey of literature should be conducted providing support for subsequent data analysis and thesis writing.

3. Dataset generation and Data Analysis (Months 5-6)

During this period, all of the aggregated data should be compiled into a single dataset. Appropriate analysis should be conducted on given data and main research conclusions should be established.

4. Thesis Writing (Months 7-8)

During this period, all of the knowledge and results obtained during the project should be compiled into the final Thesis.

Critical resources required for success of given project is access to VentureSource and VentureXpress databases. Given that academic institutions generally do have the access to given sources, obtaining access should be feasible. However, it would be critical to resolve this question as a part of initial phase of the projects (months 1-2).

Additional resources required would be software tools and hardware resources for conducting data collection and analysis. However, given that author already have access to such resources, no budget would be required.

Findings and Disseminations

We hope that given data analysis will enable us to establish positive relation between social media feedback and venture capital investments. This would enable us to formulate the model of venture capital based on publicly observable social media data that can be used in modeling future prospects of new technology ventures. Additionally, we expect this relation to be different depending on the value of "technology category" dummy variable, as it is expected for different market segments to have different sensitivity to social feedback. For example, in the case of Consumer Web Startups, this relationship is expected to be much stronger than in the case of Enterprise Startups.

References

[1] **Asur S., Huberman B. A.** – Predicting the Future with Social Media, *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*

[2] **Block J., Sandner P.** What is the effect of current financial crisis on venture capital financing? Empirical evidence from US Internet start-ups, *Venture Capital, Vol. 11, Iss 4, 2009*

[3] **Bottazzi L., Da Rin M., Hellmann T.** – The importance of Trust for Investment: Evidence from Venture Capital, *NBER Working Paper No. 16923, April 2011*

[4] **Bollen J., Mao H., Zeng X-J.** – Twitter mood predicts the stock market, *Journal of Computational Science, Vol. 2, Issue 1, March 2011*

[5] **Dolencic G.** – Determinants of the spatial reach of venture capital firms in the United States, *Master Thesis, Erasmus School of Economics, August 2010*

[6] **Giot P., Schwienbacher A.** – IPOs, trade sales and liquidations: modeling venture capital exits using survival analysis, *Journal of Banking & Finance, Vol. 31, Issue 3, March 2007*

[7] **Gompers P. A.** – Optimal Investment, Monitoring and Staging of Venture Capital, *The Journal of Finance, Vol 50, No 5.* (1995)

[8] **Gompers P., Lerner J.** – The Venture Capital Cycle, *MIT Press, second edition (2006)*

[9] **Gompers P., Kovner A., Lerner J., Scharfstein D.** – Venture Capital investment cycles : The impact of public markets, *Journal of Financial Economics 87, Vol. 87, Issue 1. January 2008*
[10] **Haeussler C., Harhoff D., Muller E.** – To be financed or not … - The Role of Patents for Venture Capital Financing, *ZEW – Centre for European Economic Research Discussion Paper No. 09-003*
[11] **Jensen B.J., Zhang M., Sobel K., Chowhudry A.** – Twitter Power: Tweets as Electronic Word of Mouth, *Journal of the American Society for Information Science and Technology, Vol. 60, Issue 11, November 2009*
[12] **Kaplan A. M., Haenlein M.** – Users of the world, unite! The challenges and opportunities of Social Media , *Business Horizons*, *Vol. 53, Issue 1. January-February 2010*
[13] **Maats F. H. E. , Metrick A., Yasuda A., Hinkes B. L. , Vershovski S.** – On the Consistency and Reliability of Venture Capital Databases, *Working Paper, Yale School of Management,* 2011
[14] **Metrick A., Yasuda A.** – Venture Capital and Other Private Equity: A Survey, *NBER Working Paper No. 16652, December 2010*
[15] **Page L., Brin S., Motwani R.** – The PageRank Citation Ranking: Bringing order to the web, *Technical Report, Stanford InfoLab, 1999*
[16] **Skeels M. M., Grudin J.** – When Social Networks Cross Boundaries: A Case Study of Workplace Use of Facebook and Linkedin, *Proceedings of the ACM 2009 international conference on Supporting group work*
[17] **Tian X.** – The causes and consequences of venture capital stage financing, *Journal of Financial Economics, Vol. 101, Issue 1, July 2011*
[18] **Teten D, Farmer C.** – Where are the Deals? Private Equity and Venture Capital Funds' Best Practices in Sourcing New Investments, *The Journal of Private Equity, Vol. 14, No. 1, Winter 2010*
[19] **Tumasjan A., Sprenger T. O., Sandner P. G. , Welpe I. M.** – Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment, *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 2010*
[20] **Wang S., Zhou H.** – Staged Financing in Venture Capital: Moral Hazard and Risks, *Journal of Corporate Finance, Vol. 10, Issue 1, January 2004*
[21] **Weiss A.** – Social Networking as a business information tool, *The Journal of Documentation, 2009/3*
[22]**Webster J. G.** – User Information Regimes: How Social Media Shape Patterns of Consumption, *Northwestern University Law Review, Spring 2010*