

Study uses text-mining to improve market intelligence on startups

13 July 2016

A researcher at The University of Texas at Arlington has created a new method that uses big data analytics and text-mining techniques to improve market intelligence and explain potential mergers and acquisitions of startup companies in the fast-moving high-technology industry.

"Industry giants like Google, Microsoft and Yahoo are spending tens of billions of dollars a year on acquiring smaller firms for market entrance, strategic intellectual property and talented employees, but face a real challenge identifying companies with the right products or technology in the vast startup universe," said Gene Moo Lee, UTA assistant professor of Information Systems and Operations Management.

"Our new approach uses big data analytics and a text-mining technique called topic modeling to identify potential matches," Lee added. "By analyzing unstructured, publicly available descriptions of any startups' [business](#), we can quantify any two firms' business, geographic, investor and social proximity and from there identify potential targets for mergers and acquisitions."

The researchers have demonstrated the applicability of their research by developing a cloud-based information system based on their method and have even launched a new company, Topic Technologies, that uses the system to offer market intelligence services on competitors, investors, acquisition targets and potential business partners to companies and startups across the high-technology sector.

Mary Whiteside, interim chair of Information Systems and Operations Management within UTA's College of Business, emphasized that this research forms an integral part of UTA's strategic focus on data-driven discovery within the Strategic Plan 2020: Bold Solutions | Global Impact.

"This research demonstrates the potential transformation [big data analytics](#) can bring to business intelligence with the use of external data sources and text mining Whiteside said. "Topic modelling provides entrepreneurs, venture capitalists and analysts a new way to navigate the constantly changing landscape of mergers and acquisitions."

For the initial analysis, the researchers used publicly available information from startup database CrunchBase on 24,382 companies, the vast majority of which were privately held, early-stage startups. For each company, they took into account the headquarter location, industry sector, cofounders, board members, key employees, investments, and the business description, which was usually limited to one or more paragraphs on the key facts about the companies' products, markets and technologies.

They then employed the text mining technique called topic modeling, which analyzes the language used in the startups' business descriptions around shared products, technologies and markets. The startups' business proximity was then quantified based on the similarity of these topic descriptions.

The likelihood of a possible merger between two companies was then computed taking into account business proximity, geographic vicinity, social links between individuals within the two firms and common investor ownership, reflecting the strongly networked nature of the startup world.

"This data-driven, analytics-based approach has proved effective in explaining mergers and acquisitions in the startup world and complements existing toolkits for measuring business proximity," Lee said. "Our system is particularly appropriate when the firms under study are small and privately held so industry classification is largely unavailable, which is the case for startups."

Lee and his co-researchers Zhan (Michael) Shi, assistant professor of Information Systems at the W.P. Carey School of Business at Arizona State University and Andrew Whinston, Hugh Cullen Chair Professor at the McCombs School of Business at the University of Texas at Austin, are also publishing their research in the leading journal *Management Information Systems Quarterly* as "Towards a Better Measure of Business Proximity: Topic Modeling for Industry Intelligence." The paper is forthcoming but currently available as a preprint.

In addition to his academic research, Lee has extensive industry experience at Samsung Electronics, AT&T Labs, Intel and Goldman Sachs. His research interests include large-scale data analytics with applications on mobile ecosystems, social network analysis and Internet security. He holds ten patents in mobile technology.

Shi's research interests reflect the interface of economics and computation, with applications in social media, online markets and innovation. His research has been published in numerous top academic journals and conferences.

Whinston is the Hugh Cullen Chair Professor in the Information, Risk and Operation Management Department at the McCombs School of Business at the University of Texas at Austin. He is also director at the Center for Research in Electronic Commerce.

Provided by University of Texas at Arlington

APA citation: Study uses text-mining to improve market intelligence on startups (2016, July 13) retrieved 22 April 2019 from <https://phys.org/news/2016-07-text-mining-intelligence-startups.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.